

Improved estimation of the MSEs and the MSE matrices for shrinkage estimators of multivariate normal means and their applications

Hisayuki Hara
Department of Geosystem Engineering
University of Tokyo

October 2007

Abstract

In this article we provide some nonnegative and positive estimators of the mean squared errors(MSEs) for shrinkage estimators of multivariate normal means. Proposed estimators are shown to improve on the uniformly minimum variance unbiased estimator(UMVUE) under a quadratic loss criterion. A similar improvement is also obtained for the estimators of the MSE matrices for shrinkage estimators. We also apply the proposed estimators of the MSE matrix to form confidence sets centered at shrinkage estimators and show their usefulness through numerical experiments.

Key words and phrases: confidence set, inadmissibility, quadratic loss, risk reduction, shrinkage estimation, UMVUE

1 Introduction.

Since Stein [15] and James and Stein [9] proved the inadmissibility of the sample mean as an estimator of a multivariate normal mean and demonstrated the superiority of the James-Stein estimator, considerable efforts have been devoted to develop the theory of the shrinkage estimation in theoretical literatures. In step with the development, shrinkage estimators have been applied to some practical problems. For example, see Fay and Herriot[7], Battese, Harter and Fuller[3] for the small-area estimation and Adkins and Eells[1] for the estimation of energy demand. As pointed out in Wan et al.[17], however, shrinkage estimation does not necessarily find many applications in empirical work. This is considered mainly because relatively little attention has been paid to evaluate the precision of shrinkage estimators. This issue is important not only for recognizing how

much it pays to use shrinkage estimators for a given problem but also for constructing confidence sets.

Joshi[10] proved the existence of confidence sets of a multivariate normal mean centered at shrinkage estimators which show higher coverage probabilities than the conventional one, a sphere centered at the observations, when the variance is known. Hwang and Casella[8] give an explicit proof that when the variance is known, a confidence set centered at the positive-part Stein estimator has uniformly higher coverage probability than the conventional one.

Carter et al.[4] discuss this problem with unknown variance in the context of linear regression models and proposed to use the estimators of the MSE and MSE matrix of the James-Stein estimator of regression coefficients to evaluate the precision of it. They derived the UMVUEs of the MSE and the MSE matrix for the James-Stein estimator and proposed to use the UMVUE of the MSE matrix to form a confidence set. They also showed that the proposed confidence set has an asymptotically smaller volume than that of the conventional one based on F statistics. Wan et al.[17] generalized their argument to a wider class of shrinkage estimators. As clarified in Section 2, however, the UMVUEs of the MSE and the MSE matrix of a shrinkage estimator have a serious drawback that they cannot always take positive or positive definite. Hence in a practical sense, the results in Carter et al.[4] and Wan et al.[17] may still not be suitable for empirical work.

Kubokawa and Srivastava[13] focused their attention on the drawback of the UMVUEs and considered the estimation of the MSEs and the MSE matrices from a decision theoretical viewpoint. They defined quadratic loss functions for evaluating estimators and proposed positive estimators of the MSE and the MSE matrix of the James-Stein estimator improving on the UMVUE under the loss functions.

In this article we take a similar approach to the one in Kubokawa and Srivastava[13] and extend their argument to the MSEs and the MSE matrices of a wider class of shrinkage estimators. We propose nonnegative and positive estimators of the MSEs and the MSE matrices of shrinkage estimators improving on the UMVUEs. We also apply the proposed estimators of the MSE matrices to form confidence sets centered at shrinkage estimators and show that they have higher coverage probabilities than the conventional one through Monte Carlo studies.

The organization of this article is as follows. In Section 2 we summarize some preliminary facts on shrinkage estimators and their MSEs and MSE matrices. Section 3 and Section 4 provide some nonnegative and positive estimators of the MSEs and the MSE matrices of shrinkage estimators improving on the UMVUEs. In Section 5.1 we give some Monte Carlo studies to confirm the theoretical results on the dominance relationship between the proposed estimators and the UMVUEs. In Section 5.2, we apply the proposed estimators of MSE matrices to form confidence sets centered at shrinkage estimators and show their usefulness through Monte Carlo studies. We end this paper with some concluding remarks in Section 6.

2 Preliminary facts on the estimation of MSE and MSE matrix for shrinkage estimators

Let \mathbf{X} and S be random variables which are independently distributed as

$$\mathbf{X} \sim N_p(\boldsymbol{\theta}, \sigma^2 \mathbf{I}_p), \quad S \sim \sigma^2 \chi_n^2, \quad (1)$$

where $N_p(\boldsymbol{\theta}, \sigma^2 \mathbf{I}_p)$ denotes p -variate normal distribution with a mean $\boldsymbol{\theta}$ and a covariance matrix $\sigma^2 \mathbf{I}_p$ and χ_n^2 denotes a chi-square variable with n degrees of freedom. Suppose that both $\boldsymbol{\theta}$ and σ^2 are unknown and that $p \geq 3$. We write $W = \|\mathbf{X}\|^2/S$. In this article we consider the following class of shrinkage estimators

$$\boldsymbol{\delta}_\phi = \left(1 - \frac{\phi(W)}{W}\right) \mathbf{X} \quad (2)$$

improving on \mathbf{X} under squared error loss. When $\phi(W) = (p-2)/(n+2)$, $\boldsymbol{\delta}_\phi$ is the James-Stein estimator (James and Stein[9])

$$\boldsymbol{\delta}_{JS} = \left(1 - \frac{p-2}{n+2} \cdot \frac{1}{W}\right) \mathbf{X}.$$

When $\phi(W) = \max(W, (p-2)/(n+2))$, $\boldsymbol{\delta}_\phi$ is the positive-part Stein estimator (Baranchik[2]),

$$\boldsymbol{\delta}_{JS}^+ = \max\left(0, 1 - \frac{p-2}{n+2} \cdot \frac{1}{W}\right) \mathbf{X}.$$

We assume that $\phi(\cdot)$ and $\phi'(\cdot)$ are finite. Most shrinkage estimators in the class (2) which have been proposed in literature satisfy the condition. The MSE matrix of $\boldsymbol{\delta}_\phi$ is defined by $\mathbf{M}(\boldsymbol{\delta}_\phi) = E[(\boldsymbol{\delta}_\phi - \boldsymbol{\theta})(\boldsymbol{\delta}_\phi - \boldsymbol{\theta})']$ and the MSE is its trace, $R(\boldsymbol{\delta}_\phi) = \text{tr} \mathbf{M}(\boldsymbol{\delta}_\phi)$. We note that since $\boldsymbol{\delta}_\phi$ improves on \mathbf{X} , $R(\boldsymbol{\delta}_\phi) \leq p\sigma^2$. In this article we consider the estimation of $R(\boldsymbol{\delta}_\phi)$ and $\mathbf{M}(\boldsymbol{\delta}_\phi)$. As Carter et al.[4] and Wan et al.[17] suggested, the estimated MSE and MSE matrix are useful as precision measures for $\boldsymbol{\delta}_\phi$. Denote by $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi)$ the UMVUE of $\mathbf{M}(\boldsymbol{\delta}_\phi)$. They proposed to use $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi)$ in a quadratic form which defines a confidence set for $\boldsymbol{\theta}$,

$$Q(\boldsymbol{\delta}_\phi) = (\boldsymbol{\delta}_\phi - \boldsymbol{\theta})' \hat{\mathbf{M}}_0^{-1}(\boldsymbol{\delta}_\phi) (\boldsymbol{\delta}_\phi - \boldsymbol{\theta}). \quad (3)$$

Let $\boldsymbol{\Gamma}$ be the orthogonal matrix such that $\boldsymbol{\Gamma} \mathbf{X} = (\|\mathbf{X}\|, 0, \dots, 0)'$ and let \mathbf{E}_{11} be the $p \times p$ matrix which has one for (1,1) element and zero elsewhere. Wan et al.[17] showed that when both $\phi(\cdot)$ and $\phi'(\cdot)$ are continuous, $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi)$ is expressed by

$$\begin{aligned} \hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi) &= \frac{S}{n} \cdot \mathbf{I}_p - S \cdot g_1(W) \cdot \mathbf{I}_p + S \cdot g_2(W) \cdot \frac{\mathbf{X} \mathbf{X}'}{\|\mathbf{X}\|^2} + S \cdot \frac{\phi^2(W)}{W} \cdot \frac{\mathbf{X} \mathbf{X}'}{\|\mathbf{X}\|^2}, \\ &= S \cdot \boldsymbol{\Gamma}' \left\{ \frac{\mathbf{I}_p}{n} - g_1(W) \cdot \mathbf{I}_p + g_3(W) \cdot \mathbf{E}_{11} \right\} \boldsymbol{\Gamma}, \end{aligned} \quad (4)$$

where

$$g_i(W) = W^{\frac{n}{2}} \left\{ \int_W^\infty t^{-\frac{n}{2}} \cdot \frac{h_i(t)}{t} dt \right\}, \quad i = 1, 2, \quad g_3(W) = g_2(W) + \frac{\phi^2(W)}{W},$$

$$h_1(W) = \frac{\phi(W)}{W}, \quad h_2(W) = 2 \left(\frac{\phi(W)}{W} - \phi'(W) \right).$$

The UMVUE of $R(\boldsymbol{\delta}_\phi)$ is obtained by

$$\hat{R}_0(\boldsymbol{\delta}_\phi) = \text{tr} \hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi) = \frac{pS}{n} - S \cdot g(W) + \frac{S\phi^2(W)}{W}, \quad (5)$$

where

$$g(W) = W^{\frac{n}{2}} \left\{ \int_W^\infty t^{-\frac{n}{2}} \cdot \frac{h(t)}{t} dt \right\}, \quad h(W) \equiv (p-2) \frac{\phi(W)}{W} + 2\phi'(W).$$

In the case of $\boldsymbol{\delta}_{JS}$, Efron and Morris[6] and Carter et al.[4] showed that $\hat{R}_0(\boldsymbol{\delta}_{JS})$ and $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS})$ are expressed by

$$\hat{R}_0(\boldsymbol{\delta}_{JS}) = \frac{pS}{n} - \left(\frac{p-2}{n+2} \right)^2 \frac{S}{W}, \quad (6)$$

$$\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}) = S \cdot \boldsymbol{\Gamma}' \left[\frac{\mathbf{I}_p}{n} - \frac{p-2}{(n+2)^2} \frac{1}{W} (2\mathbf{I}_p - (p+2)\mathbf{E}_{11}) \right] \boldsymbol{\Gamma}, \quad (7)$$

respectively. We can see that (6) and (7) coincide with (5) and (4) with $\phi(W) = (p-2)/(n+2)$. $g_1(\cdot)$ and $g_3(\cdot)$ of $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS})$ is written by

$$g_1(W) = \frac{2(p-2)}{(n+2)^2} \cdot \frac{1}{W}, \quad g_3(W) = \frac{(p-2)(p+2)}{(n+2)^2} \cdot \frac{1}{W}. \quad (8)$$

In the case of $\boldsymbol{\delta}_{JS}^+$, $\phi'(\cdot)$ is not continuous. In order to apply the argument in Wan et al.[17] to the derivation of $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$ and $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$, it requires some modifications. $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$ and $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$ are expressed as follows,

$$\hat{R}_0(\boldsymbol{\delta}_{JS}^+) = \begin{cases} -\frac{pS}{n} + SW + C_0SW^{n/2}, & W \leq \frac{p-2}{n+2}, \\ \hat{R}_0(\boldsymbol{\delta}_{JS}), & \text{otherwise,} \end{cases} \quad (9)$$

$$\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+) = \begin{cases} S \cdot \boldsymbol{\Gamma}' \left[\frac{\mathbf{I}_p}{n} - \left(\frac{2}{n} - C_1W^{n/2} \right) \mathbf{I}_p \right. \\ \quad \left. + (W + C_2W^{n/2}) \mathbf{E}_{11} \right] \boldsymbol{\Gamma}, & W \leq \frac{p-2}{n+2}, \\ \hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}), & \text{otherwise,} \end{cases} \quad (10)$$

where

$$C_0 = 2 \left(\frac{p}{n} - \frac{p-2}{n+2} \right) \cdot \left(\frac{p-2}{n+2} \right)^{-n/2},$$

$$C_1 = 2 \left(\frac{1}{n} - \frac{1}{n+2} \right) \cdot \left(\frac{p-2}{n+2} \right)^{-n/2}, \quad C_2 = \frac{4}{n+2} \left(\frac{p-2}{n+2} \right)^{-n/2}.$$

The proofs of (9) and (10) are given in the Appendix. $g_1(W)$ and $g_3(W)$ of $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$ for $W \leq (p-2)/(n+2)$ is written by

$$g_1(W) = \frac{2}{n} - C_1 W^{n/2}, \quad g_3(W) = W + C_2 W^{n/2}. \quad (11)$$

When $W > (p-2)/(n+2)$, $g_1(W)$ and $g_3(W)$ of $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$ is identical to (8).

It is easy to see from (6) and (9) that both $\hat{R}_0(\boldsymbol{\delta}_{JS})$ and $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$ have an undesirable property of taking negative values for small W . When we use an estimated MSE matrix to form a confidence set as in (3), it should be positive definite. However we can also see from (7) and (10) that both $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS})$ and $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$ are not positive definite for small W . In this article we consider the estimation of $R(\boldsymbol{\delta}_\phi)$ and $\mathbf{M}(\boldsymbol{\delta}_\phi)$ from a decision theoretical viewpoint. Kubokawa and Srivastava[13] considered the estimation of the risk reduction and the risk reduction matrix of $\boldsymbol{\delta}_{JS}$

$$R^*(\boldsymbol{\delta}_{JS}) = p\sigma^2 - R(\boldsymbol{\delta}_{JS}), \quad \mathbf{M}^*(\boldsymbol{\delta}_{JS}) = \sigma^2 \mathbf{I}_p - \mathbf{M}(\boldsymbol{\delta}_{JS}).$$

They provided a class of estimators $\hat{R}^*(\boldsymbol{\delta}_{JS})$ and $\hat{\mathbf{M}}^*(\boldsymbol{\delta}_{JS})$ improving on the UMVUEs of $R^*(\boldsymbol{\delta}_\phi)$ and $\mathbf{M}^*(\boldsymbol{\delta}_\phi)$,

$$\hat{R}_0^*(\boldsymbol{\delta}_{JS}) = \frac{pS}{n} - \hat{R}_0(\boldsymbol{\delta}_\phi), \quad \hat{\mathbf{M}}_0^*(\boldsymbol{\delta}_{JS}) = \frac{S}{n} \mathbf{I}_p - \hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi)$$

under quadratic losses and satisfying $\hat{R}(\boldsymbol{\delta}_{JS}) = (pS/n) - \hat{R}^*(\boldsymbol{\delta}_{JS}) > 0$ and $\hat{\mathbf{M}}(\boldsymbol{\delta}_{JS}) = (S/n) \mathbf{I}_p - \hat{\mathbf{M}}^*(\boldsymbol{\delta}_{JS}) > \mathbf{0}$. In the following sections we extend the argument in Kubokawa and Srivastava[13] to a wider class of shrinkage estimators $\boldsymbol{\delta}_\phi$ and propose some classes of nonnegative and positive estimators of $R(\boldsymbol{\delta}_\phi)$ and nonnegative definite and positive definite estimators of $\mathbf{M}(\boldsymbol{\delta}_\phi)$ improving on the UMVUEs quadratic loss functions.

Remark 1. *Carter et al.[4] and Wan et al.[17] considered this problem in the context of linear regression models,*

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where \mathbf{Y} is the N observations, \mathbf{A} is the $N \times p$ design matrix and $\boldsymbol{\epsilon} \sim N_N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ is the disturbance. As mentioned in Kubokawa and Srivastava[13], the setting (1) considered here can be interpreted as a canonical form of linear regression models. Let $\mathbf{B} = (\mathbf{A}'\mathbf{A})^{1/2}$ denote a $p \times p$ matrix such that $\mathbf{B}\mathbf{B}' = \mathbf{A}'\mathbf{A}$. Then \mathbf{X} , $\boldsymbol{\theta}$, S and n corresponds to $\mathbf{B}^{-1}\mathbf{A}'\mathbf{Y}$, $\mathbf{B}\boldsymbol{\beta}$, $\mathbf{Y}(\mathbf{I}_N - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\mathbf{Y}$ and $N - p$, respectively.

3 Improved estimators of the MSE

3.1 An improved nonnegative estimator of the MSE

In this section we provide a nonnegative estimator of $R(\boldsymbol{\delta}_\phi)$ improving on $\hat{R}_0(\boldsymbol{\delta}_\phi)$. For evaluating an estimator $\hat{R}(\boldsymbol{\delta}_\phi)$, we use the following quadratic loss function,

$$L(\hat{R}(\boldsymbol{\delta}_\phi) : R(\boldsymbol{\delta}_\phi)) = (\hat{R}(\boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2. \quad (12)$$

Under the loss (12), the truncated estimator $\hat{R}^{TR}(\boldsymbol{\delta}_\phi) = \max(0, \hat{R}_0(\boldsymbol{\delta}_\phi))$ obviously dominates $\hat{R}_0(\boldsymbol{\delta}_\phi)$. The estimator proposed in this section is shown to dominate $\hat{R}^{TR}(\boldsymbol{\delta}_\phi)$.

Let $a(W)$ be defined by

$$a(W) = \frac{n}{p} \left(g(W) - \frac{\phi^2(W)}{W} \right).$$

Then $\hat{R}_0(\boldsymbol{\delta}_\phi)$ is rewritten by

$$\hat{R}_0(\boldsymbol{\delta}_\phi) = pS(1 - a(W)) / n. \quad (13)$$

We introduce the following class of estimators with a function $\psi(\cdot)$,

$$\hat{R}(\psi; \boldsymbol{\delta}_\phi) = \frac{pS}{n} (1 - a(W)\psi(W)), \quad \psi(W) \leq \frac{1}{a(W)}. \quad (14)$$

Then we obtain the following theorem.

Theorem 1. *Let $\psi_0(W)$ be defined by*

$$\psi_0(W) = \max \left[\min \left(1, \frac{1}{a(W)} \right), \frac{1}{a(W)} \left(1 - \frac{n(1+W)}{n+p+2} \right) \right].$$

Then

$$\begin{aligned} \hat{R}(\psi_0; \boldsymbol{\delta}_\phi) &= \frac{pS}{n} (1 - a(W)\psi_0(W)) \\ &= \min \left[\max(\hat{R}_0(\boldsymbol{\delta}_\phi), 0), \frac{pS(1+W)}{n+p+2} \right] \end{aligned}$$

improves on $\hat{R}_0(\boldsymbol{\delta}_\phi)$ under the loss function (12).

Proof. In the proof of this theorem, we use the similar procedure which Stein[16] used to derive the improved truncated estimator of σ^2 . From the definition of $\hat{R}(\psi; \boldsymbol{\delta}_\phi)$, we have

$$\begin{aligned} &(\hat{R}(\psi; \boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \\ &= \left(\frac{pS}{n} \cdot a(W) \right)^2 \psi^2(W) - 2 \left\{ \left(\frac{pS}{n} \right)^2 a(W) - \frac{pS}{n} \cdot R(\boldsymbol{\delta}_\phi) \right\} \psi(W) \\ &\quad + \left(\frac{pS}{n} - R(\boldsymbol{\delta}_\phi) \right)^2 \end{aligned} \quad (15)$$

Let $\psi^*(W)$ be the function which minimizes $E[(\hat{R}(\psi; \boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W]$. From (15), $\psi^*(W)$ is expressed by

$$\psi^*(W) = \frac{1}{a(W)} - \frac{nE[S/\sigma^2 \mid W]}{E[(S/\sigma^2)^2 \mid W]} \cdot \frac{1}{a(W)} \cdot \frac{R(\boldsymbol{\delta}_\phi)}{p\sigma^2}. \quad (16)$$

Denote by $f_k(\cdot)$ and $f_k(\cdot; \lambda)$ the density functions of the central and the non-central χ^2 distributions with k degrees of freedom and the non-centrality parameter $\lambda = \|\boldsymbol{\theta}\|^2/\sigma^2$. Then we have

$$\begin{aligned} \frac{E[S/\sigma^2 \mid W]}{E[(S/\sigma^2)^2 \mid W]} &= \frac{\int u^2 f_p(uW; \lambda) f_n(u) du}{\int u^3 f_p(uW; \lambda) f_n(u) du} \\ &\leq \frac{\int u^2 f_p(uW) f_n(u) du}{\int u^3 f_p(uW) f_n(u) du} = \frac{1 + w}{n + p + 2}, \end{aligned} \quad (17)$$

from the monotone nondecreasingness of $f_p(uW; \lambda)/f_p(uW)$ in u . $R(\boldsymbol{\delta}_\phi) \leq p\sigma^2$. Hence from (16), (17) and the fact that $0 < R(\boldsymbol{\delta}_\phi) \leq p\sigma^2$, $\psi^*(W)$ satisfies

$$\frac{1}{a(W)} \left(1 - \frac{n(1+W)}{n+p+2} \right) \leq \psi^*(W) \leq \frac{1}{a(W)}$$

for any λ and any W such that $a(W) > 0$. Denote $\mathcal{R}^+ = \{W \mid W \geq 0\}$ and define \mathcal{A} , \mathcal{B} and \mathcal{C} as follows,

$$\begin{aligned} \mathcal{A} &= \left\{ W \mid \frac{1}{a(W)} \leq 1, a(W) > 0, W \geq 0 \right\} \\ \mathcal{B} &= \left\{ W \mid \frac{1}{a(W)} \left(1 - \frac{n(1+W)}{n+p+2} \right) \geq 1, a(W) > 0, W \geq 0 \right\} \\ \mathcal{C} &= \mathcal{R}^+ \setminus (\mathcal{A} \cup \mathcal{B}). \end{aligned}$$

We note that $E[(\hat{R}(\psi; \boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W]$ is convex on $\psi(W)$. Therefore, if we set

$$\psi(W) = \min \left(1, \frac{1}{a(W)} \right) \quad \text{for } W \in \mathcal{A},$$

$\psi^*(W)$ satisfies $\psi^*(W) \leq \psi(W) \leq 1$ for $W \in \mathcal{A}$, which implies

$$E[(\hat{R}(\psi; \boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W \in \mathcal{A}] \leq E[(\hat{R}_0(\boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W \in \mathcal{A}]. \quad (18)$$

Similarly, if we set

$$\psi(W) = \max \left[\frac{1}{a(W)} \left(1 - \frac{n(1+W)}{n+p+2} \right), 1 \right] \quad \text{for } W \in \mathcal{B},$$

$\psi^*(W)$ satisfies $\psi^*(W) \geq \psi(W) \geq 1$ for $W \in \mathcal{B}$ and then

$$\mathbb{E}[(\hat{R}(\psi; \boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W \in \mathcal{B}] \leq \mathbb{E}[(\hat{R}_0(\boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W \in \mathcal{B}]. \quad (19)$$

For $W \in \mathcal{C}$, if we set $\psi(W) = 1$,

$$\mathbb{E}[(\hat{R}(\psi; \boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W \in \mathcal{C}] = \mathbb{E}[(\hat{R}_0(\boldsymbol{\delta}_\phi) - R(\boldsymbol{\delta}_\phi))^2 \mid W \in \mathcal{C}]. \quad (20)$$

By combining (18), (19) and (20), we can complete the proof. \square

We have shown that $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ dominates $\hat{R}_0(\boldsymbol{\delta}_\phi)$ under the loss (12). We note that the unbiasedness of $\hat{R}_0(\boldsymbol{\delta}_\phi)$ was not used in the proof. By following the proof, we can see that any estimator in the class (14) which does not satisfy

$$0 \leq \hat{R}(\psi; \boldsymbol{\delta}_\phi) \leq \frac{pS(1+W)}{n+p+2} \quad (21)$$

is improved by the one which is truncated it to satisfy (21). In other words, (21) is a necessary condition that $\hat{R}(\psi; \boldsymbol{\delta}_\phi)$ is admissible in the class (14).

In the case of $\boldsymbol{\delta}_{JS}$, $a(W) = n(p-2)^2/(p(n+2)^2W)$ from (6). Then \mathcal{B} for $\boldsymbol{\delta}_{JS}$ is written by

$$\mathcal{B} = \left\{ W \left| \left(\frac{n+2}{p-2} \right)^2 \cdot \frac{pW}{n} \cdot \left(1 - \frac{n(1+W)}{n+p+2} \right) \geq 1, W \geq 0 \right. \right\}.$$

Hence if (p, n) satisfy

$$\left(\frac{p+2}{n} \right)^2 - \frac{4(n+p+2)}{p} \left(\frac{p-2}{n+2} \right)^2 \leq 0,$$

$\mathcal{B} = \emptyset$ and then $\hat{R}(\psi_0; \boldsymbol{\delta}_{JS})$ is identical to $\hat{R}^{TR}(\boldsymbol{\delta}_{JS})$. In general, when n and p are large, $\mathcal{B} = \emptyset$ and hence $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ is identical to $\hat{R}^{TR}(\boldsymbol{\delta}_\phi)$. However if $\mathcal{B} \neq \emptyset$, $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ is shown to improve on $\hat{R}^{TR}(\boldsymbol{\delta}_\phi)$.

Corollary 1. *If (p, n) satisfies $\mathcal{B} \neq \emptyset$, $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ improves on $\hat{R}^{TR}(\boldsymbol{\delta}_\phi)$ under the loss (12).*

Proof. As mentioned above, $\hat{R}^{TR}(\boldsymbol{\delta}_\phi)$ is identical to $\hat{R}(\psi; \boldsymbol{\delta}_\phi)$ with

$$\psi(W) = \begin{cases} \min \left(1, \frac{1}{a(W)} \right), & \text{if } W \in \mathcal{A}, \\ 1, & \text{otherwise.} \end{cases}$$

By following the proof of Theorem 1, we can see that $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ improves on $\hat{R}^{TR}(\boldsymbol{\delta}_\phi)$ under the loss (12) when $\mathcal{B} \neq \emptyset$. \square

We present in Figure 1 the shape of $1 - a(W)\psi_0(W)$ for $\boldsymbol{\delta}_{JS}$ and $\boldsymbol{\delta}_{JS}^+$ with $(p, n) = (5, 5)$. The dotted line and the solid line represent $1 - a(W)$ and $1 - a(W)\psi_0(W)$, respectively. We can see that when $\hat{R}_0(\boldsymbol{\delta}_\phi) < 0$, $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ shrinks towards 0 and even when $\hat{R}_0(\boldsymbol{\delta}_\phi) > 0$, $\hat{R}(\psi_0; \boldsymbol{\delta}_\phi)$ slightly shrinks it towards 0.

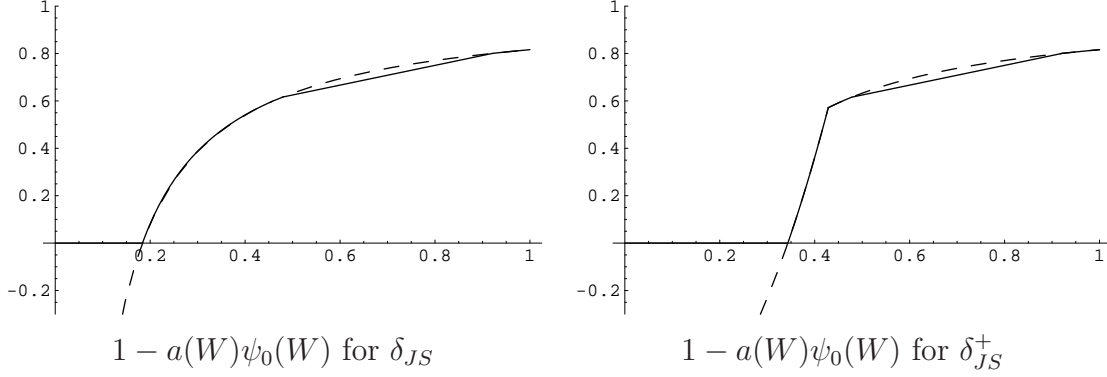


Figure 1: $1 - a(W)\psi_0(W)$ of δ_{JS} and δ_{JS}^+ when $(p, n) = (5, 5)$

3.2 Improved estimation of the risk reduction and positive estimation of MSE

In the previous section we have shown that $\hat{R}(\psi_0; \delta_\phi)$ improves on $\hat{R}_0(\delta_\phi)$. However we can easily see that $\hat{R}(\psi_0; \delta_\phi)$ can take zero with positive probability. From a practical viewpoint, $\hat{R}(\psi_0; \delta_\phi)$ may still be undesirable. In order to obtain positive estimators, we take an similar approach to the one in Kubokawa and Srivastava[13] and consider the improved estimation of the risk reduction $R^*(\delta_\phi) = p\sigma^2 - R(\delta_\phi)$. In accordance with Kubokawa and Srivastava[13], for evaluating the estimator $\hat{R}^*(\delta_\phi)$, we use the following quadratic loss function,

$$L(\hat{R}^*(\delta_\phi); R^*(\delta_\phi)) = (\hat{R}^*(\delta_\phi) - R^*(\delta_\phi))^2. \quad (22)$$

From (13), the UMVUE of $R^*(\delta_\phi)$ is $\hat{R}_0^*(\delta_\phi) = pSa(W)/n$. In order to improve on $\hat{R}_0^*(\delta_\phi)$, we introduce the class of estimators $\hat{R}^*(\psi; \delta_\phi) = pSa(W)\psi(W)/n$. We can see that $pS/n - \hat{R}^*(\psi; \delta_\phi)$ is identical to $\hat{R}(\psi; \delta_\phi)$ defined in (14). It is well known that $R^*(\delta_\phi)/\sigma^2$ depends only on λ (e.g. Efron and Morris[6]). We suppose that $R^*(\delta_\phi)/\sigma^2$ is maximized at $\lambda = 0$. If $\phi(W)/W$ is nonincreasing, δ_ϕ satisfy this condition (e.g. Kubokawa[11], Casella[5]). Many shrinkage estimators including δ_{JS} and δ_{JS}^+ satisfy this condition. Denote $R^*(\delta_\phi)/\sigma^2$ with $\lambda = 0$ by $\alpha_{p,n}(\delta_\phi)$. By using the similar argument in Theorem 1, we can obtain the following result.

Theorem 2. Let $W_{p,n}(\delta_\phi)$ be the solution of the equation

$$\frac{1+W}{a(W)} = \frac{p(n+p+2)}{n} \cdot \frac{1}{\alpha_{p,n}(\delta_\phi)}. \quad (23)$$

Define $\psi_1(W)$ by

$$\psi_1(W) = \max \left(1, \frac{n}{p} \cdot \frac{1+W_{p,n}(\delta_\phi)}{n+p+2} \cdot \frac{\alpha_{p,n}(\delta_\phi)}{a(W)} \right).$$

Then $\hat{R}^*(\psi_1; \delta_\phi)$ improves on $\hat{R}_0^*(\delta_\phi)$.

The proof is similar to the proof of Theorem 1 and is given in the Appendix.

By using the similar argument in Kubokawa and Srivastava[13], we can also obtain the following result.

Theorem 3. *Suppose that*

- (i) $a(W)$ is nonincreasing ;
- (ii) $\psi(W)$ is nondecreasing and $\lim_{W \rightarrow \infty} \psi(W) = 1$;
- (iii) $\psi(W) \geq \psi_2(W)$ where

$$\psi_2(W) = \min \left(1, \frac{\alpha_{p,n}(\boldsymbol{\delta}_\phi)}{pa(W)} \right).$$

Then $\hat{R}^*(\psi; \boldsymbol{\delta}_\phi)$ improves on $\hat{R}_0^*(\boldsymbol{\delta}_\phi)$ under the loss (22). Especially, $\hat{R}^*(\psi_2; \boldsymbol{\delta}_\phi)$ improves on $\hat{R}_0^*(\boldsymbol{\delta}_\phi)$.

The proof of this theorem is also given in the Appendix. We note that we can see from (6) and (9) that $\boldsymbol{\delta}_{JS}$ and $\boldsymbol{\delta}_{JS}^+$ satisfies the condition (i). The results of Theorem 2 and Theorem 3 lead us to propose the following estimators of $\hat{R}(\boldsymbol{\delta}_\phi)$,

$$\begin{aligned} \hat{R}(\psi_1; \boldsymbol{\delta}_\phi) &= \frac{pS}{n} - \hat{R}^*(\psi_1; \boldsymbol{\delta}_\phi) \\ &= \max \left[\frac{pS}{n}(1 - a(W)), \frac{pS}{n} \left(1 - \frac{n(1 + W_{p,n}(\boldsymbol{\delta}_\phi))}{n + p + 2} \cdot \frac{\alpha_{p,n}(\boldsymbol{\delta}_\phi)}{p} \right) \right], \\ \hat{R}(\psi_2; \boldsymbol{\delta}_\phi) &= \frac{pS}{n} - \hat{R}^*(\psi_2; \boldsymbol{\delta}_\phi) \\ &= \max \left[\frac{pS}{n}(1 - a(W)), \frac{pS}{n} \left(1 - \frac{\alpha_{p,n}(\boldsymbol{\delta}_\phi)}{p} \right) \right]. \end{aligned}$$

Since $\alpha_{p,n}(\boldsymbol{\delta}_\phi) < p$, $\hat{R}(\psi_2; \boldsymbol{\delta}_\phi)$ is always positive. If p and n satisfy

$$\gamma_{p,n}(\boldsymbol{\delta}_\phi) := \frac{n(1 + W_{p,n}(\boldsymbol{\delta}_\phi))}{n + p + 2} < \frac{p}{\alpha_{p,n}(\boldsymbol{\delta}_\phi)(\boldsymbol{\delta}_\phi)}, \quad (24)$$

$\hat{R}(\psi_1; \boldsymbol{\delta}_\phi) > 0$ with probability one. We can prove $\gamma_{p,n}(\boldsymbol{\delta}_{JS}) < p/\alpha_{p,n}(\boldsymbol{\delta}_\phi)(\boldsymbol{\delta}_{JS})$ analytically. The proof is given in the Appendix.

We present $\gamma_{p,n}(\boldsymbol{\delta}_{JS})$, $\gamma_{p,n}(\boldsymbol{\delta}_{JS}^+)$, $W_{p,n}(\boldsymbol{\delta}_{JS})$ and $W_{p,n}(\boldsymbol{\delta}_{JS}^+)$ for $(p, n) = (5, 5)$, $(10, 5)$, $(5, 10)$ and $(10, 10)$ in Table 1 and 2. We can see $\gamma_{p,n}(\boldsymbol{\delta}_{JS}) < 1$ and $\gamma_{p,n}(\boldsymbol{\delta}_{JS}^+) < 1$ for all cases. Since $p/\alpha_{p,n}(\boldsymbol{\delta}_\phi)(\boldsymbol{\delta}_\phi) > 1$, $\gamma_{p,n}(\boldsymbol{\delta}_{JS})$ and $\gamma_{p,n}(\boldsymbol{\delta}_{JS}^+)$ satisfy (24). Many shrinkage estimators seem to satisfy (24). However it seems to be difficult to prove (24) for general $\boldsymbol{\delta}_\phi$ analytically. We note that $\gamma_{p,n}(\boldsymbol{\delta}_{JS}) < 1$ and $\gamma_{p,n}(\boldsymbol{\delta}_{JS}^+) < 1$ mean that $\psi_1(\cdot)$ does not satisfy the conditions in Theorem 3. Hence from Table 1, $\hat{R}(\psi_1; \boldsymbol{\delta}_\phi)$ is not necessarily included in the class of estimators in Theorem 3.

Table 1: $\gamma_{p,n}(\boldsymbol{\delta}_{JS})$ and $\gamma_{p,n}(\boldsymbol{\delta}_{JS}^+)$ for some p and n

(p, n)	(5, 5)	(10, 5)	(5, 10)	(10, 10)
$\gamma_{p,n}(\boldsymbol{\delta}_{JS})$	0.6795	0.7452	0.7774	0.8228
$\gamma_{p,n}(\boldsymbol{\delta}_{JS}^+)$	0.6399	0.7056	0.7484	0.7921

Table 2: $W_{p,n}(\boldsymbol{\delta}_{JS})$ and $W_{p,n}(\boldsymbol{\delta}_{JS}^+)$ for some p and n

(p, n)	(5, 5)	(10, 5)	(5, 10)	(10, 10)
$W_{p,n}(\boldsymbol{\delta}_{JS})$	0.6307	1.533	0.3216	0.8102
$W_{p,n}(\boldsymbol{\delta}_{JS}^+)$	0.5357	1.399	0.2722	0.7427

So far we considered the improved estimation of $R^*(\boldsymbol{\delta}_\phi)$. Next we consider to evaluate the resulting estimators $\hat{R}(\psi_1; \boldsymbol{\delta}_\phi)$ and $\hat{R}(\psi_2; \boldsymbol{\delta}_\phi)$ as estimators of $R(\boldsymbol{\delta}_\phi)$. Based on the fact that (21) is a necessary condition that $\hat{R}(\psi; \boldsymbol{\delta}_\phi)$ is admissible in the class (14), we can provide estimators improving on $\hat{R}(\psi_1; \boldsymbol{\delta}_\phi)$ and $\hat{R}(\psi_2; \boldsymbol{\delta}_\phi)$ under the loss (12).

Theorem 4. Suppose that $\mathcal{B} \neq \emptyset$. Let $\psi_1^{TR}(W)$ and $\psi_2^{TR}(W)$ be defined by

$$\psi_j^{TR}(W) = \max \left[\psi_j(W), \frac{1}{a(W)} \left(1 - \frac{n(1+W)}{n+p+2} \right) \right].$$

Then $\hat{R}(\psi_j^{TR}; \boldsymbol{\delta}_\phi)$ improves on $\hat{R}(\psi_j; \boldsymbol{\delta}_\phi)$ under the loss (12).

Based on the proof of Theorem 1, the proof is easy and omitted. It is clear that $\hat{R}(\psi_j^{TR}; \boldsymbol{\delta}_\phi) > 0$ when $\hat{R}(\psi_j; \boldsymbol{\delta}_\phi) > 0$. The dominance relation between $\hat{R}(\psi_j^{TR}; \boldsymbol{\delta}_\phi)$ and $\hat{R}_0(\boldsymbol{\delta}_\phi)$ is interesting. But it seems difficult to clarify it at this point.

4 Improved estimators of the MSE matrix

4.1 An improved nonnegative definite estimation of the MSE matrix

In this section we provide a nonnegative definite estimator of $\boldsymbol{M}(\boldsymbol{\delta}_\phi)$ improving on $\hat{\boldsymbol{M}}_0(\boldsymbol{\delta}_\phi)$. For evaluating an estimator $\hat{\boldsymbol{M}}(\boldsymbol{\delta}_\phi)$, we use the following squared loss functions,

$$L(\hat{\boldsymbol{M}}(\boldsymbol{\delta}_\phi); \boldsymbol{M}(\boldsymbol{\delta}_\phi)) = \text{tr}(\hat{\boldsymbol{M}}(\boldsymbol{\delta}_\phi) - \boldsymbol{M}(\boldsymbol{\delta}_\phi))^2. \quad (25)$$

We note that $L(\hat{\boldsymbol{M}}(\boldsymbol{\delta}_\phi); \boldsymbol{M}(\boldsymbol{\delta}_\phi))$ is the sum of quadratic losses of each element. In order to consider the improvement of $\hat{\boldsymbol{M}}_0(\boldsymbol{\delta}_\phi)$, we introduce the following class of estimators

with functions $\xi(\cdot)$ and $\eta(\cdot)$,

$$\begin{aligned}\hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_\phi) &= S \cdot \boldsymbol{\Gamma}' \left\{ \frac{\mathbf{I}_p}{n} - g_1(W) \xi(W) \cdot (\mathbf{I}_p - \mathbf{E}_{11}) \right. \\ &\quad \left. + (g_3(W) - g_1(W) \eta(W)) \cdot \mathbf{E}_{11} \right\} \boldsymbol{\Gamma} \\ &= S \boldsymbol{\Gamma}' \mathbf{L}(\xi, \eta; \boldsymbol{\delta}_\phi) \boldsymbol{\Gamma},\end{aligned}\tag{26}$$

where

$$\mathbf{L}(\xi, \eta; \boldsymbol{\delta}_\phi) = \text{diag}(l_0(\eta; \boldsymbol{\delta}_\phi), l_1(\xi; \boldsymbol{\delta}_\phi), \dots, l_1(\xi; \boldsymbol{\delta}_\phi)),$$

$$l_0(\eta; \boldsymbol{\delta}_\phi) = \frac{1}{n} - g_1(W) \eta(W) + g_3(W), \quad l_1(\xi; \boldsymbol{\delta}_\phi) = \frac{1}{n} - g_1(W) \xi(W).$$

It is clear that $\hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_\phi)$ with $\xi(W) = 1$ and $\eta(W) = 1$ is identical to $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi)$. $\hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_\phi)$ with $\eta(W) = 1$ coincides with the class which Kubokawa and Srivastava[13] considered for the James-Stein estimator. In the case of positive-part Stein estimator, however, $g_3(W) - g_1(W)$ is written by

$$g_3(W) - g_1(W) = \begin{cases} \left(\frac{2}{n+2} + \frac{2}{n} \right) W^{n/2} \left(\frac{p-2}{n+2} \right)^{-n/2} + W - \frac{2}{n}, & W \leq \frac{p-2}{n+2}, \\ \frac{p(p-2)}{(n+2)^2} \cdot \frac{1}{W}, & \text{otherwise.} \end{cases}$$

Denote $l_0(\eta; \boldsymbol{\delta}_{JS}^+)$ with $\eta(W) = 1$ by $l_0(\boldsymbol{\delta}_{JS}^+)$. Then we can see that $l_0(\boldsymbol{\delta}_{JS}^+) < 0$ for small W . This means that any estimator $\hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_{JS}^+)$ with $\eta(W) = 1$ does not always take nonnegative definite. So we consider to find improved estimators from the class (26).

We assume that $g_1(W)$ is nonincreasing. From (8) and (11), we can see that $\boldsymbol{\delta}_{JS}$ and $\boldsymbol{\delta}_{JS}^+$ satisfy this condition. Let u and v be random variables which distribute as chi-square distribution with $p+2j$ and n degrees of freedom, respectively. Then define $\beta_{p,n}^{(1)}(\boldsymbol{\delta}_\phi)$

$$\beta_{p,n}^{(1)}(\boldsymbol{\delta}_\phi) := \inf_{j \geq 0} \beta_{p,n}^{(1)}(j; \boldsymbol{\delta}_\phi),$$

where

$$\begin{aligned}\beta_{p,n}^{(1)}(j; \boldsymbol{\delta}_\phi) &= \mathbb{E} \left[\frac{2(p-1)\phi(u/v)}{u/v} - \frac{(p+2j-1)b(u/v)}{p+2j} \right], \\ b(W) &= \frac{4\phi(W)}{W} + \frac{(n+2)\phi^2(W)}{W} - 4\phi'(W) - 4\phi(W)\phi'(W).\end{aligned}$$

Figure 2 represents the behavior of $\beta_{p,n}^{(1)}(j; \boldsymbol{\delta}_\phi)$ of $\boldsymbol{\delta}_{JS}$ and $\boldsymbol{\delta}_{JS}^+$ with $(p, n) = (5, 5), (10, 10)$ computed by Monte Carlo simulation. The result of Figure 2 indicates that $\beta_{p,n}^{(1)}(\boldsymbol{\delta}_{JS}) \geq 0$

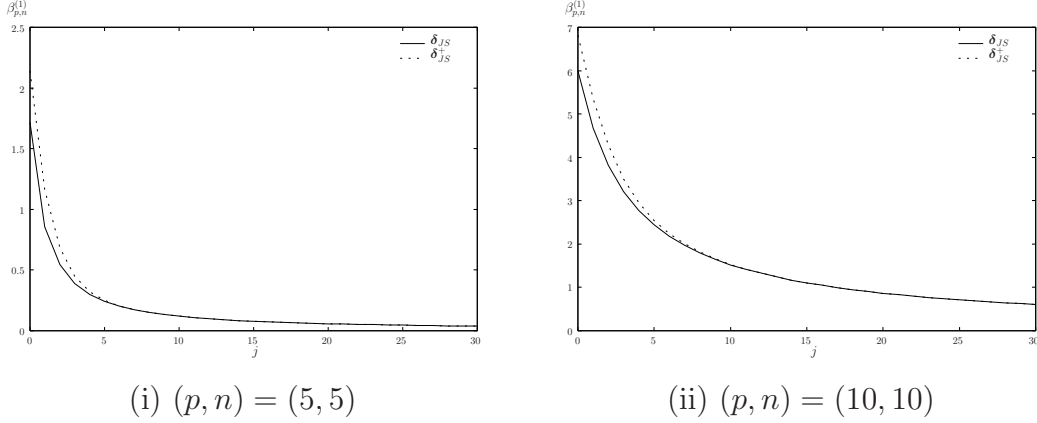


Figure 2: $\beta_{p,n}^{(1)}(j; \delta_\phi)$ of δ_{JS} and δ_{JS}^+ with $(p, n) = (5, 5), (10, 10)$

and $\beta_{p,n}^{(1)}(\delta_{JS}^+) \geq 0$. We obtained the same results numerically with respect to other (p, n) and other δ_ϕ .

By using the similar argument in the proof of Theorem 1, we can obtain the following theorem.

Theorem 5. Suppose that $\beta_{p,n}^{(1)}(\delta_\phi) \geq 0$. Let $\xi_0(W)$ and $\eta_0(W)$ be defined by

$$\begin{aligned} \xi_0(W) &= \max \left[\min \left(1, \frac{1}{ng_1(W)} \right), \frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{1+W}{n+p+1} \right) \right], \\ \eta_0(W) &= \min \left[1, \frac{1}{g_1(W)} \left(\frac{1}{n} + g_3(W) \right) \right] \end{aligned}$$

Then $\hat{\mathbf{M}}(\xi_0, \eta_0; \delta_\phi)$ improves on $\hat{\mathbf{M}}_0(\delta_\phi)$ under the loss (25).

The proof is similar to the one of Theorem 1 and is given in the Appendix. By following the proof of the theorem and by using the same argument in Section 3.1, we can see that

$$\frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{1+W}{n+p+2} \right) \leq \xi(W) \leq \frac{1}{ng_1(W)}, \quad (27)$$

is a necessary condition on $\hat{\mathbf{M}}(\xi, \eta; \delta_\phi)$ to be admissible in the class (26).

We can easily see that $\hat{\mathbf{M}}(\xi_0, \eta_0; \delta_\phi)$ is always nonnegative definite. However $l_1(\xi_0; \delta_\phi) = 0$ when $\xi_0(W) = 1/(ng_1(W))$ and $l_0(\eta_0; \delta_\phi) = 0$ when $\eta_0(W) = (1/n + g_3(W))/g_1(W)$, i.e. $\hat{\mathbf{M}}(\xi_0, \eta_0; \delta_\phi)$ is not always positive definite. As mentioned in the previous section, from a practical viewpoint, the inverse matrices of estimators of the MSE matrix are required to form a confidence set as (3). In this sense $\hat{\mathbf{M}}(\xi_0, \eta_0; \delta_\phi)$ may still be undesirable. In the next section we provide some positive definite estimators by considering the improvement of the UMVUE of the risk reduction matrix $\mathbf{M}^*(\delta_\phi)$.

4.2 An improved estimation of the risk reduction matrix and positive-definite estimation of the MSE matrix

Following the argument in Kubokawa and Srivastava[13], we consider the estimation of MSE reduction matrices $\mathbf{M}^*(\boldsymbol{\delta}_\phi) = \sigma^2 \mathbf{I}_p - \mathbf{M}(\boldsymbol{\delta}_\phi)$. As a criterion, we use the following quadratic loss function,

$$L(\hat{\mathbf{M}}^*(\boldsymbol{\delta}_\phi); \mathbf{M}^*(\boldsymbol{\delta}_\phi)) = \text{tr}(\hat{\mathbf{M}}^*(\boldsymbol{\delta}_\phi) - \mathbf{M}^*(\boldsymbol{\delta}_\phi))^2. \quad (28)$$

We consider the following class derived from (26),

$$\begin{aligned} \hat{\mathbf{M}}^*(\xi, \eta; \boldsymbol{\delta}_\phi) &= \frac{S}{n} \mathbf{I}_p - \hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_\phi) \\ &= S \cdot \boldsymbol{\Gamma}' \{ g_1(W) \xi(W) \cdot (\mathbf{I}_p - \mathbf{E}_{11}) - (g_3(W) - g_1(W) \eta(W)) \cdot \mathbf{E}_{11} \} \boldsymbol{\Gamma} \\ &= S \cdot \boldsymbol{\Gamma}' \mathbf{L}^*(\xi, \eta; \boldsymbol{\delta}_\phi) \boldsymbol{\Gamma}, \end{aligned} \quad (29)$$

where

$$\begin{aligned} \mathbf{L}^*(\xi, \eta; \boldsymbol{\delta}_\phi) &= \text{diag}(l_0^*(\eta; \boldsymbol{\delta}_\phi), l_1^*(\xi; \boldsymbol{\delta}_\phi), \dots, l_1^*(\xi; \boldsymbol{\delta}_\phi)), \\ l_0^*(\eta; \boldsymbol{\delta}_\phi) &= g_1(W) \eta(W) - g_3(W) = \frac{1}{n} - l_0(\eta; \boldsymbol{\delta}_\phi), \\ l_1^*(\xi; \boldsymbol{\delta}_\phi) &= g_1(W) \xi(W) = \frac{1}{n} - l_1(\xi; \boldsymbol{\delta}_\phi) \end{aligned}$$

Let u , v and $b(W)$ be defined as in the previous section. Define $\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)$ by

$$\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) := \sup_{j \geq 0} \beta_{p,n}^{(2)}(j; \boldsymbol{\delta}_\phi),$$

where

$$\beta_{p,n}^{(2)}(j; \boldsymbol{\delta}_\phi) = \mathbb{E} \left[\frac{2\phi(u/v)}{u/v} - \frac{b(u/v)}{p+2j} \right].$$

Figure 3 represents the behavior of $\beta_{p,n}^{(2)}(j; \boldsymbol{\delta}_{JS})$ and $\beta_{p,n}^{(2)}(j; \boldsymbol{\delta}_{JS}^+)$ with $(p, n) = (5, 5)$, $(10, 10)$ computed by Monte Carlo simulation. Table 3 presents $\beta_{p,n}^{(2)}(\boldsymbol{\delta}_{JS})$ and $\beta_{p,n}^{(2)}(\boldsymbol{\delta}_{JS}^+)$ when (p, n) is $(5, 5)$, $(5, 10)$, $(10, 5)$ and $(10, 10)$. Kubokawa and Srivastava[13] proved that $\sup_{j \geq 0} \beta_{p,n}^{(2)}(j; \boldsymbol{\delta}_{JS})$ are attained at $j = 0, 1$. They also showed numerically that $\sup_{j \geq 0} \beta_{p,n}^{(2)}(j; \boldsymbol{\delta}_{JS}^+)$ are attained at $j = 0$. The result of Figure 3 is consistent to their results.

Similar to the argument in Theorem 5, we can obtain the following result.

Theorem 6. *Let the solutions of the equations*

$$\frac{1+W}{n+p+2} \cdot \frac{\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{g_1(W)} = 1, \quad \frac{1}{g_1(W)} \left(g_3(W) + \frac{1+W}{n+p+2} \cdot \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right) = 1$$

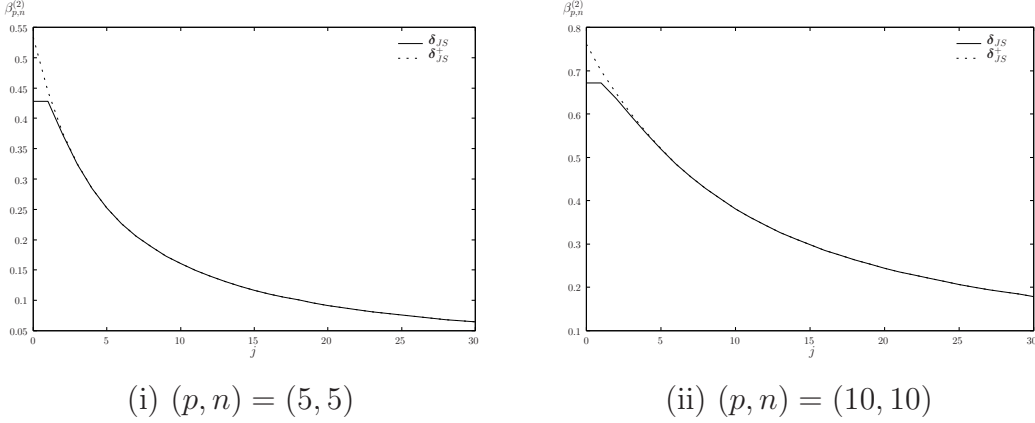


Figure 3: $\beta_{p,n}^{(2)}(j : \delta_\phi)$ of δ_{JS} and δ_{JS}^+ with $(p, n) = (5, 5)$ and $(10, 10)$

Table 3: $\beta_{p,n}^{(2)}(\delta_{JS})$ and $\beta_{p,n}^{(2)}(\delta_{JS}^+)$ for some p and n

(p, n)	$(5, 5)$	$(10, 5)$	$(5, 10)$	$(10, 10)$
$\beta_{p,n}^{(2)}(\delta_{JS})$	0.4260	0.5704	0.5008	0.6718
$\beta_{p,n}^{(2)}(\delta_{JS}^+)$	0.5332	0.6611	0.5173	0.7616

be denoted by $W_{p,n}^\xi(\delta_\phi)$ and $W_{p,n}^\eta(\delta_\phi)$, respectively. Define $\xi_1(W)$ and $\eta_1(W)$ by

$$\xi_1(W) = \min \left(1, \frac{1 + W_{n,p}^\xi(\delta_\phi)}{n + p + 2} \cdot \frac{\beta_{p,n}^{(2)}(\delta_\phi)}{g_1(W)} \right),$$

$$\eta_1(W) = \min \left[1, \frac{1}{g_1(W)} \left(g_3(W) + \frac{1 + W_{n,p}^\eta(\delta_\phi)}{n + p + 2} \cdot \beta_{p,n}^{(2)}(\delta_\phi) \right) \right].$$

Then $\hat{\mathbf{M}}^*(\xi_1, \eta_1; \delta_\phi)$ improves on $\hat{\mathbf{M}}_0^*(\delta_\phi)$ under the loss (28).

The proof of this theorem is given in the Appendix. By using the similar procedure to Kubokawa and Srivastava[13], we can also obtain the following theorem.

Theorem 7. Suppose that

- (i) $g_1(W)$ is nonincreasing and $g_3(W)/g_1(W)$ is nondecreasing ;
- (ii) $\xi(\cdot)$ and $\eta(\cdot)$ are nondecreasing and satisfy

$$\lim_{W \rightarrow \infty} \xi(W) = 1, \quad \lim_{W \rightarrow \infty} \eta(W) = 1;$$

- (iii) $\xi(W) \geq \xi_2(W)$ and $\eta(W) \geq \eta_2(W)$ where

$$\xi_2(W) = \min \left(1, \frac{1}{g_1(W)} \cdot \frac{\beta_{p,n}^{(2)}(\delta_\phi)}{n + 2} \right),$$

$$\eta_2(W) = \min \left(1, \frac{1}{g_1(W)} \left(g_3(W) + \frac{\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{n+2} \right) \right).$$

Then $\hat{\mathbf{M}}^*(\xi, \eta; \boldsymbol{\delta}_\phi)$ improves on $\hat{\mathbf{M}}_0^*(\boldsymbol{\delta}_\phi)$ under the loss (28).

The proof of this theorem is also given in the Appendix. We can easily see from (8) and (11) that the condition (i) is satisfied for $\boldsymbol{\delta}_{JS}$ and $\boldsymbol{\delta}_{JS}^+$. We note that if $g_3(W) - g_1(W) \geq 0$ for all $W \geq 0$, $\eta_1(W) = 1$ and $\eta_2(W) = 1$ with probability one.

The results of Theorem 6 and Theorem 7 lead us to propose the estimators of the MSE matrix $\mathbf{M}(\boldsymbol{\delta}_\phi)$ as

$$\hat{\mathbf{M}}(\xi_i, \eta_i; \boldsymbol{\delta}_\phi) = \frac{S}{n} \mathbf{I}_p - \hat{\mathbf{M}}^*(\xi_i, \eta_i; \boldsymbol{\delta}_\phi), \quad i = 1, 2.$$

If (p, n) satisfies

$$\begin{aligned} \gamma_{p,n}^\xi(\boldsymbol{\delta}_\phi) &:= \frac{n(1 + W_{p,n}^\xi(\boldsymbol{\delta}_\phi))\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{n + p + 2} \leq 1, \\ \gamma_{p,n}^\eta(\boldsymbol{\delta}_\phi) &:= \frac{n(1 + W_{p,n}^\eta(\boldsymbol{\delta}_\phi))\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{n + p + 2} \leq 1, \end{aligned}$$

$\hat{\mathbf{M}}(\xi_1, \eta_1; \boldsymbol{\delta}_\phi)$ is always positive definite and if (p, n) satisfies

$$\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)/(n+2) \leq 1/n,$$

$\hat{\mathbf{M}}(\xi_2, \eta_2; \boldsymbol{\delta}_\phi)$ is always positive. Table 4 present $\gamma_{p,n}^\xi(\boldsymbol{\delta}_{JS})$, $\gamma_{p,n}^\xi(\boldsymbol{\delta}_{JS}^+)$ and $\gamma_{p,n}^\eta(\boldsymbol{\delta}_{JS}^+)$ when (p, n) is (5, 5), (5, 10), (10, 5) and (10, 10). From Table 3 and 4, we can see that the above conditions are satisfied for (p, n) in the tables. In Table 5 we also presents $W_{p,n}^\xi(\boldsymbol{\delta}_{JS})$, $W_{p,n}^\xi(\boldsymbol{\delta}_{JS}^+)$ and $W_{p,n}^\eta(\boldsymbol{\delta}_{JS}^+)$ for the same (p, n) as in Table 4.

Recall the fact that (27) is a necessary condition on $\hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_\phi)$ to be admissible in the class (26). By using the similar argument in Theorem 3, we can also obtain the estimators improving on $\hat{\mathbf{M}}(\xi_i, \eta_i; \boldsymbol{\delta}_\phi)$, $i = 1, 2$ under the loss (25).

Table 4: $\gamma_{p,n}^\xi(\boldsymbol{\delta}_{JS})$, $\gamma_{p,n}^\xi(\boldsymbol{\delta}_{JS}^+)$ and $\gamma_{p,n}^\eta(\boldsymbol{\delta}_{JS}^+)$

(p, n)	(5, 5)	(10, 5)	(5, 10)	(10, 10)
$\gamma_{p,n}^\xi(\boldsymbol{\delta}_{JS})$	0.4312	0.6143	0.5273	0.7548
$\gamma_{p,n}^\xi(\boldsymbol{\delta}_{JS}^+)$	0.4963	0.6690	0.6102	0.8170
$\gamma_{p,n}^\eta(\boldsymbol{\delta}_{JS}^+)$	0.2708	0.2567	0.4130	0.4014

Table 5: $W_{p,n}^\xi(\boldsymbol{\delta}_{JS})$, $W_{p,n}^\xi(\boldsymbol{\delta}_{JS}^+)$ and $W_{p,n}^\eta(\boldsymbol{\delta}_{JS}^+)$

(p, n)	(5, 5)	(10, 5)	(5, 10)	(10, 10)
$W_{p,n}^\xi(\boldsymbol{\delta}_{JS})$	1.4198	2.6577	0.7901	1.472
$W_{p,n}^\xi(\boldsymbol{\delta}_{JS}^+)$	1.2336	2.4405	0.6829	1.360
$W_{p,n}^\eta(\boldsymbol{\delta}_{JS}^+)$	0.2185	0.3202	0.1391	0.1596

Theorem 8. Suppose that $\beta_{p,n}^{(1)}(\boldsymbol{\delta}_\phi) \geq 0$. Let $\xi_i^{TR}(W)$ for $i = 1, 2$ be defined by

$$\xi_i^{TR}(W) = \max \left[\xi_i(W), \frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{1+W}{n+p+1} \right) \right]$$

Then if

$$\left\{ W \mid \xi_i^{TR}(W) = \frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{1+W}{n+p+1} \right) \right\} \neq \emptyset,$$

$\hat{\mathbf{M}}(\xi_i^{TR}, \eta_i; \boldsymbol{\delta}_\phi)$ improves on $\hat{\mathbf{M}}(\xi_i, \eta_i; \boldsymbol{\delta}_\phi)$ under the loss (25).

The proof of this theorem is easy and omitted. It is obvious that when $\hat{\mathbf{M}}(\xi_i, \eta_i; \boldsymbol{\delta}_\phi)$ is positive definite, $\hat{\mathbf{M}}(\xi_i^{TR}, \eta_i; \boldsymbol{\delta}_\phi)$ is also positive definite. The dominance relation between $\hat{\mathbf{M}}(\xi_j^{TR}, \eta_j; \boldsymbol{\delta}_\phi)$ and $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_\phi)$ under the loss (25) is interesting. In the same way as the case of the MSE, however, it seems difficult to clarify it at this point.

5 Monte Carlo studies

5.1 Risk performance of proposed estimators

In this section we study the risk performance of the proposed estimators of the MSE and the MSE matrix for the positive part Stein estimator through Monte Carlo studies with 100,000 replications. Figure 4 represents the performance of risks $E[L(\hat{R}(\boldsymbol{\delta}_{JS}^+); R(\boldsymbol{\delta}_{JS}^+))]$ of $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$, $\hat{R}(\psi_0; \boldsymbol{\delta}_{JS}^+)$, $\hat{R}(\psi_1^{TR}; \boldsymbol{\delta}_{JS}^+)$ and $\hat{R}(\psi_2^{TR}; \boldsymbol{\delta}_{JS}^+)$. Figure 5 represents the performance of risks $E[L(\hat{\mathbf{M}}(\boldsymbol{\delta}_{JS}^+); \mathbf{M}(\boldsymbol{\delta}_{JS}^+))]$ of $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$, $\hat{\mathbf{M}}(\psi_0; \boldsymbol{\delta}_{JS}^+)$, $\hat{\mathbf{M}}(\psi_1^{TR}; \boldsymbol{\delta}_{JS}^+)$ and $\hat{\mathbf{M}}(\psi_2^{TR}; \boldsymbol{\delta}_{JS}^+)$. We set $\boldsymbol{\theta}$, σ^2 , λ , p and n as

- $\boldsymbol{\theta} = (\sqrt{\lambda/p}, \dots, \sqrt{\lambda/p})'$;
- $\sigma^2 = 1$;
- $0 \leq \lambda \leq 30$;
- $(p, n) = (5, 5), (10, 5), (5, 10)$ and $(10, 10)$;

The summary of experiments is as follows.

- We can see from the figures that every proposed estimator uniformly improves on the UMVUE.
- In Section 3, we have shown that $\hat{R}(\psi_0; \boldsymbol{\delta}_{JS}^+)$ improves on $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$ under the loss (12). However we could not give theoretical proofs that $\hat{R}(\psi_1^{TR}; \boldsymbol{\delta}_{JS}^+)$ and $\hat{R}(\psi_2^{TR}; \boldsymbol{\delta}_{JS}^+)$ improve on $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$ under the loss (12) as estimators of the MSE. However we can see from Figure 4 that $\hat{R}(\psi_1^{TR}; \boldsymbol{\delta}_{JS}^+)$ and $\hat{R}(\psi_2^{TR}; \boldsymbol{\delta}_{JS}^+)$ seem to uniformly improve on $\hat{R}_0(\boldsymbol{\delta}_{JS}^+)$ under the loss (12).

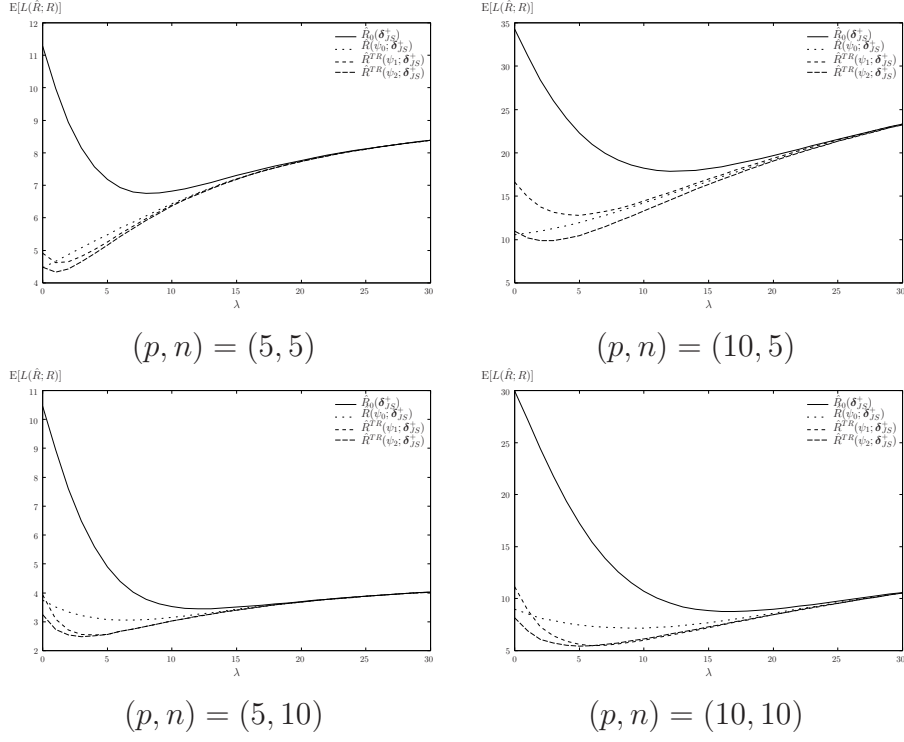


Figure 4: Estimated risks of the UMVUE and proposed estimators of $R(\delta_{JS}^+)$

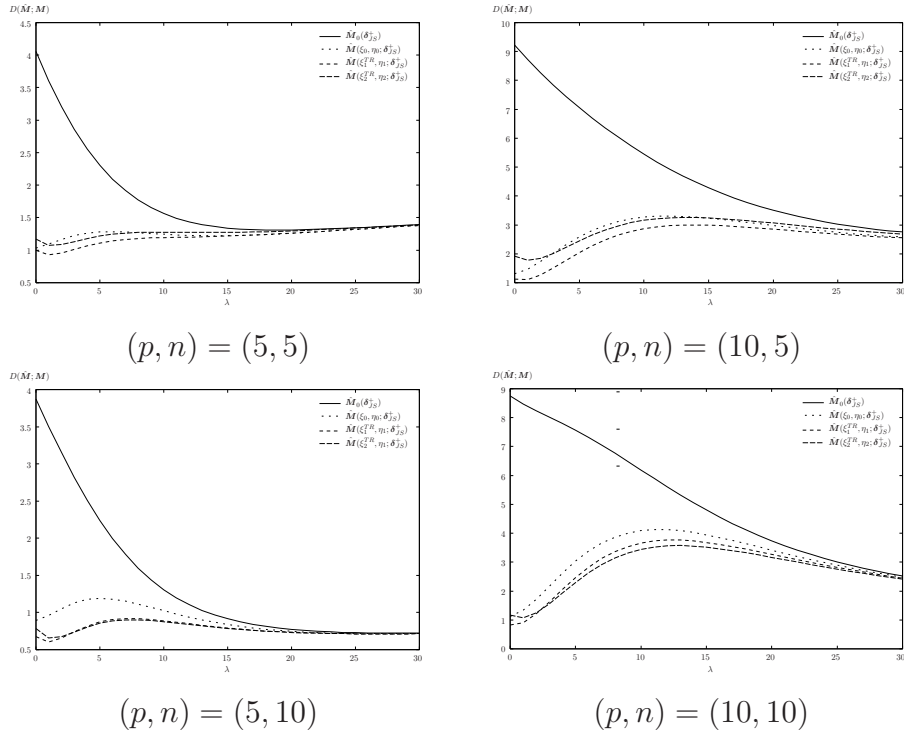


Figure 5: Estimated risks for the UMVUE and proposed estimators of $M(\delta_{JS}^+)$

- Also in the case of estimators of the MSE matrix, we can see from Figure 5 that $\hat{\mathbf{M}}(\psi_1^{TR}; \boldsymbol{\delta}_{JS}^+)$ and $\hat{\mathbf{M}}(\psi_2^{TR}; \boldsymbol{\delta}_{JS}^+)$ which are derived from improved estimators of $\mathbf{M}^*(\boldsymbol{\delta}_{JS}^+)$ seem to improve on $\hat{\mathbf{M}}_0(\boldsymbol{\delta}_{JS}^+)$ under the loss (25).
- When λ is small, proposed estimators show large rate of risk reduction. Conversely, as λ gets larger, the rate of risk reduction of the proposed estimators gets smaller.

5.2 Confidence sets centered at shrinkage estimators

In this section we study the performance of confidence sets (3) centered at the positive part Stein estimator. Let $c = F_{0.95}(p, n)$ be 95% percentile of F distribution with degrees of freedom p and n . Then the conventional 95% confidence set of $\boldsymbol{\theta}$ is written by

$$C_0 : Q_0 = \frac{(\mathbf{X} - \boldsymbol{\theta})'(\mathbf{X} - \boldsymbol{\theta})/p}{S/n} \leq c.$$

In what follows, we write $\hat{\mathbf{M}}_1 = \hat{\mathbf{M}}(\xi_1^{TR}, \eta_1; \boldsymbol{\delta}_{JS}^+)$ and $\hat{\mathbf{M}}_2 = \hat{\mathbf{M}}(\xi_2^{TR}, \eta_2; \boldsymbol{\delta}_{JS}^+)$ for notational simplicity. Following Carter et al.[4] and Wan et al.[17], we first consider the following three alternatives to C_0 ,

$$\begin{aligned} C_1 : Q_1 &= \frac{(\boldsymbol{\delta}_{JS}^+ - \boldsymbol{\theta})' \hat{\mathbf{M}}_1^{-1} (\boldsymbol{\delta}_{JS}^+ - \boldsymbol{\theta})}{p} \leq c, \\ C_2 : Q_2 &= \frac{(\boldsymbol{\delta}_{JS}^+ - \boldsymbol{\theta})' \hat{\mathbf{M}}_2^{-1} (\boldsymbol{\delta}_{JS}^+ - \boldsymbol{\theta})}{p} \leq c, \\ C_3 : Q_3 &= \frac{(\boldsymbol{\delta}_{JS}^+ - \boldsymbol{\theta})' (\boldsymbol{\delta}_{JS}^+ - \boldsymbol{\theta})/p}{S/n} \leq c. \end{aligned}$$

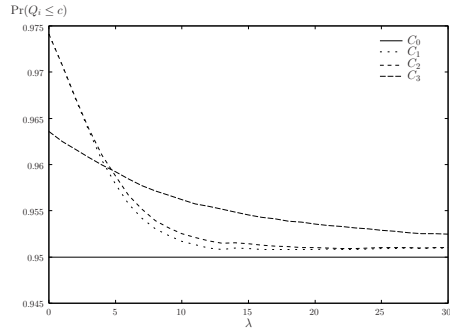
Let $Vol(C_i)$ denote the volume of C_i , i.e.

$$\begin{aligned} Vol(C_0) &= Vol(C_3) = \frac{(S/n)^{p/2} \cdot (c\pi)^{p/2}}{\Gamma(p/2 + 1)}, \\ Vol(C_1) &= \frac{|\hat{\mathbf{M}}_1|^{1/2} (c\pi)^{p/2}}{\Gamma(p/2 + 1)}, \quad Vol(C_2) = \frac{|\hat{\mathbf{M}}_2|^{1/2} (c\pi)^{p/2}}{\Gamma(p/2 + 1)}. \end{aligned}$$

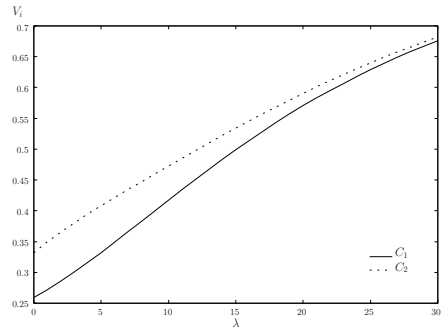
Define V_1 and V_2 by

$$V_1 = \frac{E[Vol(C_1)]}{E[Vol(C_0)]} = \frac{E[Vol(C_1)]}{E[Vol(C_3)]}, \quad V_2 = \frac{E[Vol(C_2)]}{E[Vol(C_0)]} = \frac{E[Vol(C_2)]}{E[Vol(C_3)]}.$$

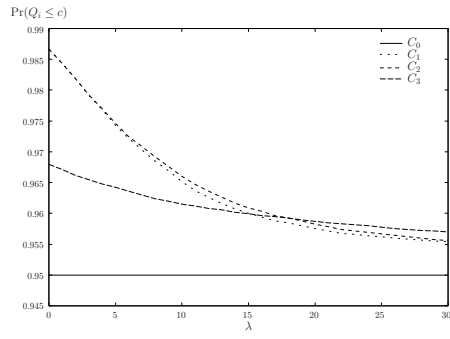
Figure 6 represent coverage probabilities of C_0 to C_3 , V_1 and V_2 computed through Monte Carlo method with 100,000 replications. We set $\boldsymbol{\theta}$, σ^2 , λ , p and n as the experiments in the previous section. From the figures, we can see that when λ is small, the coverage probabilities of C_1 to C_3 are larger than that of C_0 . When (p, n) is (5, 5) and (10, 5), the coverage probabilities of C_1 to C_3 are larger than 95% for $0 \leq \lambda \leq 30$. When (p, n) is



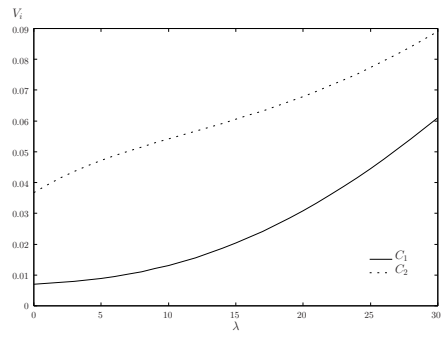
$\Pr(Q_i < c)$ for $(p, n) = (5, 5)$



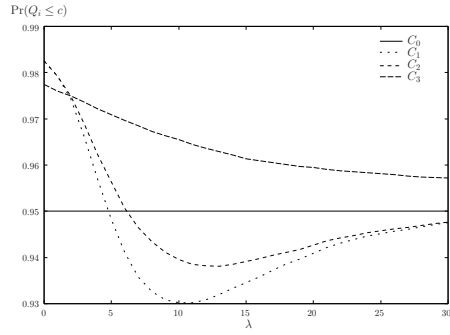
V_1 and V_2 for $(p, n) = (5, 5)$



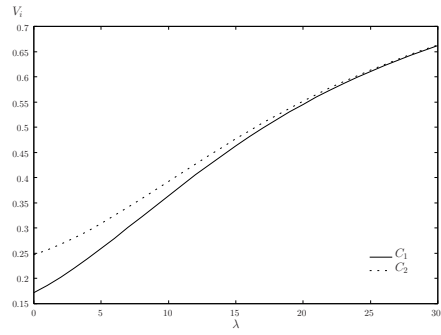
$\Pr(Q_i < c)$ for $(p, n) = (10, 5)$



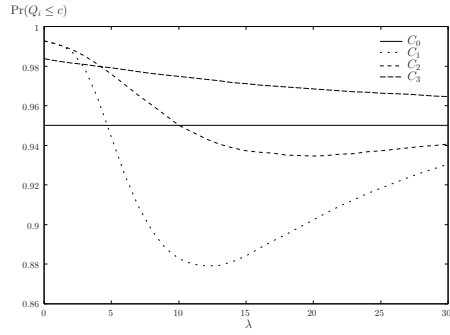
V_1 and V_2 for $(p, n) = (10, 5)$



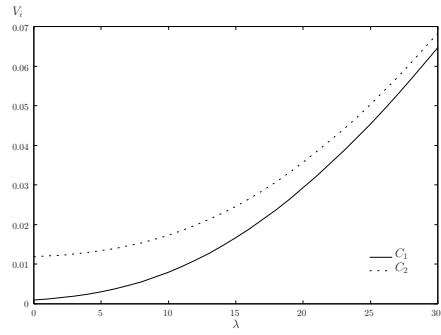
$\Pr(Q_i < c)$ for $(p, n) = (5, 10)$



V_1 and V_2 for $(p, n) = (5, 10)$



$\Pr(Q_i < c)$ for $(p, n) = (10, 10)$



V_1 and V_2 for $(p, n) = (10, 10)$

Figure 6: Coverage probabilities of Q_0 to Q_3 and V_1 and V_2 .

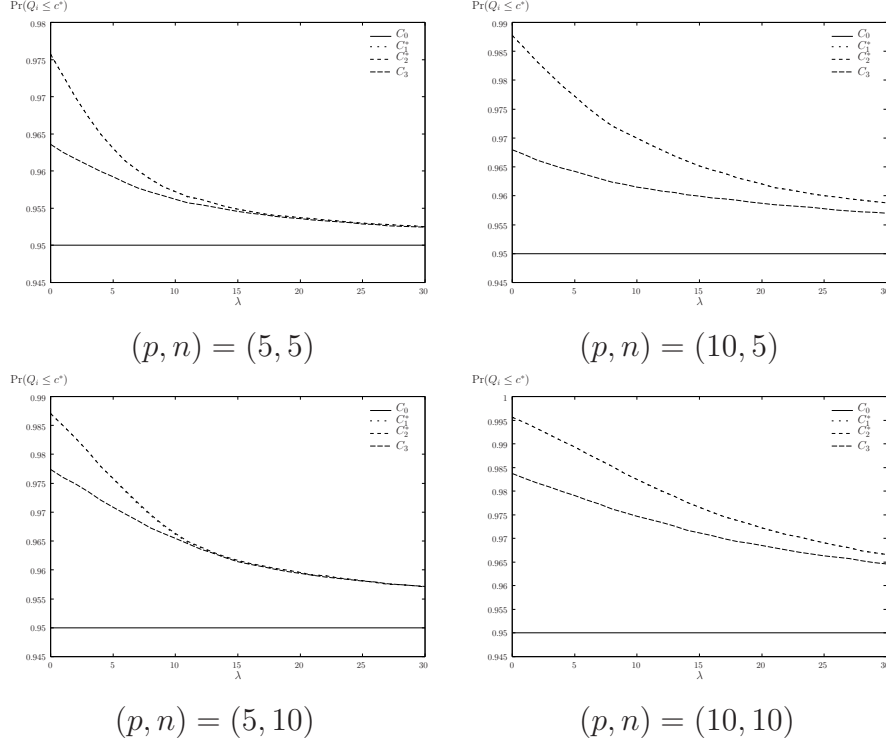


Figure 7: Coverage probabilities of C_1^* and C_2^*

(5, 10) and (10, 10), however, the coverage probabilities of C_1 and C_2 drop to less than 95% as λ gets larger. The expected volumes of C_1 and C_2 are far smaller than those of C_0 and C_3 . These results suggest that when λ and n is small, C_1 and C_2 seem to have higher coverage probabilities and smaller volumes than C_0 . On the other hand, while the expected volume of C_3 is larger than that of C_1 and C_2 , the coverage probabilities of C_3 is always larger than 95%. In this sense, C_3 may be the most practical among C_0 to C_3 .

So we next consider the following confidence sets,

$$C_1^* : Q_1^* = \frac{(\delta_{JS}^+ - \theta)' \hat{\mathbf{M}}_1^{-1} (\delta_{JS}^+ - \theta)}{p} \leq \frac{(S/n) \cdot c}{|\hat{\mathbf{M}}_1|^{1/p}},$$

$$C_2^* : Q_2^* = \frac{(\delta_{JS}^+ - \theta)' \hat{\mathbf{M}}_2^{-1} (\delta_{JS}^+ - \theta)}{p} \leq \frac{(S/n) \cdot c}{|\hat{\mathbf{M}}_2|^{1/p}}.$$

Then we have $Vol(C_1^*) = Vol(C_2^*) = Vol(C_0) = Vol(C_3)$. Denote $c^* = (S/n) \cdot c / |\hat{\mathbf{M}}_1|^{1/p}$. Figure 7 represents the behavior of the coverage probabilities of C_0 , C_1^* , C_2^* and C_3 . The behavior of the coverage probabilities of C_1^* and C_2^* are almost the same and two dashed lines which represent the behavior of them overlap each other and look like just one line. We can see from Figure 7 that the coverage probability of C_1^* and C_2^* are larger than that of C_0 and larger or at least the same level as that of C_3 . Especially when λ is small, the coverage probability of C_1^* and C_2^* are considerably larger than that of C_0 and C_3 . Since the volumes of C_1^* and C_2^* are equal to C_0 and C_3 , we can say that C_1^* and C_2^* are useful

for practical use.

6 Concluding remarks

In this article we considered the estimation of the MSE and the MSE matrix of shrinkage estimators from a decision theoretical viewpoint. We generalize the argument of Carter et al.[4], Wan et al.[17] and Kubokawa and Srivastava[13] and provided some nonnegative and positive estimators improving on the UMVUE. We also proposed some confidence sets by using the positive definite estimator of the MSE matrices and showed through Monte Carlo studies that proposed confidence sets have equal volume and attain higher coverage probabilities than the conventional one. These results suggest that the proposed estimators are useful for evaluating the precision of shrinkage estimators.

Carter et al.[4] and Wan et al.[17] derived asymptotic coverage probabilities of their proposed confidence region. By using their argument, we may provide better confidence region.

The argument in this article may also be available to the estimation of MSE and MSE matrix in the small-area problem discussed in Prasad and Rao[14]. Since the small-area estimation is an illustrative example of the application of shrinkage estimators(e.g. Fay and Herriot[7]), it would be interesting if we could extend the results here to the problem. We leave these problem to our future research.

Appendix

A The derivation of (9) and (10)

We first derive (9). Following Efron and Morris(1976),

$$\begin{aligned} R(\delta_\phi) &= \sigma^2 \mathbb{E} \left[p - \left(2(p-2) \frac{\phi(W)}{W} - (n+2) \frac{\phi^2(W)}{W} + 4\phi'(W) + 4\phi(W)\phi'(W) \right) \right] \\ &= \sigma^2 \mathbb{E} \left[p - \left(2(p-2) \frac{\phi(W)}{W} + 4\phi'(W) \right) \right] + \mathbb{E} \left[S \frac{\phi^2(W)}{W} \right]. \end{aligned}$$

The UMVU estimator of $\sigma^2 \mathbb{E}[p]$ is pS/n . With respect to the term

$$\sigma^2 \mathbb{E} \left[2(p-2) \frac{\phi(W)}{W} + 4\phi'(W) \right],$$

we use the χ^2 identity with an absolutely continuous function $g(W)$ by Efron and Morris(1976),

$$\mathbb{E}[g(W)S] = \sigma^2 \mathbb{E}[ng(W) - 2g'(W)W].$$

Let $g(W)$ be a solution of the following differential equation of first order,

$$ng(W) - 2g'(W)W = 2h(W),$$

where

$$h(W) \equiv (p-2)\frac{\phi(W)}{W} + 2\phi'(W).$$

Then Wan et al.[17] showed that the UMVUE of $\hat{R}_0(\boldsymbol{\delta}_\phi)$ can be written as in (5). The general solution of $g(W)$ is written by

$$g(W) = W^{\frac{n}{2}} \left\{ \int_W^\infty t^{-\frac{n}{2}} \cdot \frac{h(t)}{t} dt + C_0 \right\}, \quad (30)$$

where C_0 is a constant. We note that $g(W)$ is required to be absolutely continuous. When $h(W)$ is continuous, $g(W)$ is also continuous. From the assumption $\phi(W) < \infty$ and $\phi'(W) < \infty$, $E[g(W)]$ exists if and only if $C_0 = 0$. However in the case of $\boldsymbol{\delta}_{JS}^+$, $h(W)$ is not continuous and $g(W)$ with $C_0 = 0$ is also not continuous. We can easily show that $g(W)$ is absolutely continuous and $E[g(W)] < \infty$ if and only if C_0 satisfies

$$C_0 = \begin{cases} 2 \left(\frac{p}{n} - \frac{p-2}{n+2} \right) \cdot \left(\frac{p-2}{n+2} \right)^{-n/2}, & \text{if } W \leq \frac{p-2}{n+2} \\ 0, & \text{otherwise.} \end{cases}$$

By following Wan et al.[17] and the above argument, the derivation of (10) is similar and omitted. \square

B Proof of Theorem 2

The proof is similar to the proof of Theorem 1. Let $\psi^*(W)$ be the function which minimizes $E[(\hat{R}^*(\psi; \boldsymbol{\delta}_\phi) - R^*(\boldsymbol{\delta}_\phi))^2 | W]$. Then

$$\begin{aligned} \psi_\lambda^*(W) &= \frac{n}{p} \cdot \frac{E[S/\sigma^2 | W]}{E[(S/\sigma^2)^2 | W]} \cdot \frac{R^*(\boldsymbol{\delta}_\phi)}{\sigma^2} \cdot \frac{1}{a(W)} \\ &\leq \frac{n}{p} \cdot \frac{1+W}{n+p+2} \cdot \frac{\alpha_{p,n}(\boldsymbol{\delta}_\phi)}{a(W)} \\ &\leq \frac{n}{p} \cdot \frac{1+W_{p,n}(\boldsymbol{\delta}_\phi)}{n+p+2} \cdot \frac{\alpha_{p,n}(\boldsymbol{\delta}_\phi)}{a(W)}. \end{aligned}$$

By using the similar argument in the proof of Theorem 1, we can show that $\hat{R}^*(\psi_1; \boldsymbol{\delta}_\phi)$ dominates $\hat{R}_0^*(\boldsymbol{\delta}_\phi)$. \square

C Proof of $\hat{R}^*(\psi_1; \boldsymbol{\delta}_{JS}) > 0$

Since $\alpha_{p,n}(\boldsymbol{\delta}_{JS}) < p$, it suffices to show that $W_{p,n}(\boldsymbol{\delta}_{JS}) < (p+2)/n$. In the case of $\boldsymbol{\delta}_{JS}$, $a(W) = n(p-2)^2/(p(n+2)^2W)$ from (6) and $\alpha_{p,n}(\boldsymbol{\delta}_{JS}) = (n(p-2))/(n+2)$ (e.g. Kubokawa and Srivastava[13]). Hence (23) is equivalent to

$$\zeta(W) := W(1+W) = \frac{n+p+2}{n} \cdot \frac{p-2}{n+2}.$$

Since $\zeta(W)$ is increasing for $W \geq 0$ and

$$\zeta\left(\frac{p+2}{n}\right) = \frac{p+2}{n} \cdot \frac{n+p+2}{n} \geq \frac{n+p+2}{n} \cdot \frac{p-2}{n+2},$$

we have $W_{p,n}(\boldsymbol{\delta}_{JS}) < (p+2)/n$. □

D Proof of Theorem 3

For an estimator $\hat{R}^*(\delta_\phi)$, denote $D(\hat{R}^*(\delta_\phi)) = \mathbb{E}[L(\hat{R}^*(\boldsymbol{\delta}_\phi); R^*(\boldsymbol{\delta}_\phi))]$. It suffices to show that when $\psi(\cdot)$ satisfies the condition of the theorem, $D(\hat{R}_0^*(\delta_\phi)) - D(\hat{R}^*(\psi; \delta_\phi)) \geq 0$. By using the same argument as the proof of Theorem 3 in Kubokawa and Srivastava[13], we have

$$\begin{aligned} & D(\hat{R}_0^*(\delta_\phi)) - D(\hat{R}^*(\psi; \delta_\phi)) \\ &= \mathbb{E} \left[\int_1^\infty \frac{d}{dt} \left\{ \left(\frac{pS}{n} \right)^2 a^2(W) \psi^2(tW) - 2 \left(\frac{pS}{n} \right) a(W) \psi(tW) R^*(\boldsymbol{\delta}_\phi) \right\} dt \right] \\ &\geq \mathbb{E} \left[\int_1^\infty \left\{ 2 \left(\frac{pS}{n} \right)^2 a^2(W) \psi'(tW) - \left(\frac{pS}{n} \right) a(W) \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right\} W \psi'(W) dt \right] \\ &= \frac{2p}{n} \times \int \int \int_0^{vw} \left\{ \frac{p}{n} \cdot va^2(x/v) \psi(w) - a(x/v) \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right\} \\ &\quad \times \psi'(w) f_p(x; \lambda) f_n(v) dx dv dw \\ &= \int \mathbb{E}^v \left[\int_0^{vw} g_1(x/v) f_p(x; \lambda) dx \right. \\ &\quad \left. \times \left\{ \frac{p}{n} \cdot \frac{\int_0^{vw} a^2(x/v) f_p(x; \lambda) dx}{\int_0^{vw} a(x/v) f_p(x; \lambda) dx} \cdot v \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right\} \right] \phi'(w) dw \\ &\geq \frac{2p}{n} \int \mathbb{E}^v \left[\int_0^{vw} a(x/v) f_p(x; \lambda) dz \cdot \left\{ \frac{p}{n} \cdot va(w) \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right\} \right] \phi'(w) dw \\ &\geq \inf_w \left\{ \frac{2p}{n} \cdot \mathbb{E}^v \left[\int_0^{vw} a(x/v) f_p(x; \lambda) dz \cdot \left\{ \frac{p}{n} \cdot va(w) \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right\} \right] \phi'(w) \right\}, \quad (31) \end{aligned}$$

where $\mathbb{E}^v[\cdot]$ denotes the expectation with respect to $v \sim \chi_p^2$. The first inequality follows from the fact that $\alpha_{p,n}(\boldsymbol{\delta}_\phi) \geq R^*(\boldsymbol{\delta}_\phi)$. The second equality follows from the transformation, $(x, v, w) = (WS, S, tW)$ with $dt dW dS = x^{-1} dx dv dw$. The second inequality follows from the nonincreasingness of $a(\cdot)$. Since $a(\cdot)$ is nonincreasing, $\int_0^{vw} a(x/v) f_p(x; \lambda) dz$ is non-decreasing in v . With respect to $q(v) = (p/n)va(w)\phi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi)$, there exists v_0 such

that $q(v) \leq 0$ for $0 < v \leq v_0$ and $q(v) > 0$ for $v > v_0$. Then we have from Kubokawa[12]

$$\begin{aligned} \mathbb{E}^v \left[\int_0^{vw} a(x/v) f_p(x; \lambda) dz \cdot \left\{ \frac{p}{n} \cdot va(w) \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right\} \right] \\ \geq \int_0^{v_0 w} a(x/v) f_p(x; \lambda) dz \cdot \mathbb{E}^v \left[\frac{p}{n} \cdot va(w) \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right]. \end{aligned}$$

If $\psi(\cdot)$ satisfy the condition of the theorem,

$$\mathbb{E}^v \left[\frac{p}{n} \cdot va(w) \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \right] = pa(w) \psi(w) - \alpha_{p,n}(\boldsymbol{\delta}_\phi) \geq 0.$$

Therefore we can complete the proof. \square

E Proof of Theorem 5

Define $\mathbf{M}_I(\boldsymbol{\delta}_\phi) = \sigma^{-2} \mathbf{M}(\boldsymbol{\delta}_\phi)$ and $\mathbf{M}_I^*(\boldsymbol{\delta}_\phi) = \mathbf{I}_p - \sigma^{-2} \mathbf{M}(\boldsymbol{\delta}_\phi)$. Before we give a proof, we present a lemma required to prove the theorem.

Lemma 1. $\beta_{p,n}^{(1)}$ satisfies

$$\beta_{p,n}^{(1)}(\boldsymbol{\delta}_\phi) \leq \text{tr} \left(\mathbf{I}_p - \frac{\mathbf{X}\mathbf{X}'}{\|\mathbf{X}\|^2} \right) \cdot \mathbf{M}_I^*(\boldsymbol{\delta}_\phi).$$

By following Section 2.2 in Kubokawa and Srivastava[13], the proof of this lemma is easy and omitted here.

Proof of Theorem 5. Let $\xi^*(\cdot)$ and $\eta^*(\cdot)$ be the functions $\xi(\cdot)$ and $\eta(\cdot)$ which minimize $\mathbb{E}[\text{tr}(\hat{\mathbf{M}}(\xi, \eta; \boldsymbol{\delta}_\phi) - \mathbf{M}(\boldsymbol{\delta}_\phi))^2 \mid W]$. Similar to (16), we have

$$\begin{aligned} \xi^*(W) &= \frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{1}{\mathbb{E}[(S/\sigma^2)^2 \mid W]} \cdot \text{tr} \left\{ \mathbb{E} \left[\frac{S}{\sigma^2} \left(\mathbf{I}_p - \frac{\mathbf{X}\mathbf{X}'}{\|\mathbf{X}\|^2} \right) \cdot \frac{\mathbf{M}_I(\boldsymbol{\delta}_\phi)}{p-1} \mid W \right] \right\} \right), \\ &= \frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{\mathbb{E}[S/\sigma^2 \mid W]}{\mathbb{E}[(S/\sigma^2)^2 \mid W]} \right. \\ &\quad \left. + \frac{1}{\mathbb{E}[(S/\sigma^2)^2 \mid W]} \cdot \mathbb{E} \left[\frac{S}{\sigma^2} \text{tr} \left(\mathbf{I}_p - \frac{\mathbf{X}\mathbf{X}'}{\|\mathbf{X}\|^2} \right) \cdot \frac{\mathbf{M}_I^*(\boldsymbol{\delta}_\phi)}{p-1} \mid W \right] \right), \end{aligned}$$

$$\begin{aligned} \eta^*(W) &= \frac{1}{g_1(W)} \left(\frac{1}{n} + g_2(W) + \frac{\phi^2(W)}{W} \right. \\ &\quad \left. - \frac{1}{\mathbb{E}[(S/\sigma^2)^2 \mid W]} \cdot \text{tr} \left\{ \mathbb{E} \left[\frac{S}{\sigma^2} \cdot \frac{\mathbf{X}\mathbf{X}'}{\|\mathbf{X}\|^2} \cdot \mathbf{M}_I(\boldsymbol{\delta}_\phi) \mid W \right] \right\} \right). \end{aligned}$$

We note that $\xi^*(\cdot)$ does not depend on $\eta(\cdot)$ and that $\eta^*(\cdot)$ does not depend on $\xi(\cdot)$. From (17), Lemma 1, the assumption that $\beta_{p,n}^{(1)}(\boldsymbol{\delta}_\phi) \geq 0$ and the facts that

$$\mathbf{I}_p - \frac{\mathbf{X}\mathbf{X}'}{W} = \mathbf{\Gamma}'(\mathbf{I}_p - \mathbf{E}_{11})\mathbf{\Gamma} \geq \mathbf{0}, \quad \frac{\mathbf{X}\mathbf{X}'}{W} \geq \mathbf{0}, \quad M_I(\boldsymbol{\delta}_\phi) \geq \mathbf{0}. \quad (32)$$

we have

$$\begin{aligned} \frac{1}{g_1(W)} \left(\frac{1}{n} - \frac{1+W}{n+p+2} \right) &\leq \psi^*(W) \leq \frac{1}{ng_1(W)}, \\ \eta^*(W) &\leq \frac{1}{g_1(W)} \left(\frac{1}{n} + g_3(W) \right). \end{aligned}$$

By using the same argument as in the proof of Theorem 1, we can complete the proof. \square

F Proof of Theorem 6

We present a lemma required to prove the theorem.

Lemma 2. $\beta_{p,n}^{(2)}$ satisfies

$$\begin{aligned} \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) &\geq \text{tr} \frac{\mathbf{X}\mathbf{X}'}{\|\mathbf{X}\|^2} \cdot \mathbf{M}_I^*(\boldsymbol{\delta}_\phi), \\ (p-1)\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) &\geq \text{tr} \left(\mathbf{I}_p - \frac{\mathbf{X}\mathbf{X}'}{\|\mathbf{X}\|^2} \right) \cdot \mathbf{M}_I^*(\boldsymbol{\delta}_\phi). \end{aligned}$$

By following Kubokawa and Srivastava[13], the proof of this lemma is also easy and omitted here.

Proof of Theorem 6. The proof is similar to the one of Theorem 5. Let $\xi^*(\cdot)$ and $\eta^*(\cdot)$ be the functions $\xi(\cdot)$ and $\eta(\cdot)$ which minimize $\mathbb{E}[\text{tr}(\hat{\mathbf{M}}^*(\xi, \eta; \boldsymbol{\delta}_\phi) - \mathbf{M}^*(\boldsymbol{\delta}_\phi))^2 \mid W]$. Then we have

$$\xi^*(W) = \frac{1}{g_1(W)} \cdot \frac{1}{\mathbb{E}[(S/\sigma^2)^2 \mid W]} \cdot \text{tr} \left\{ \mathbb{E} \left[\frac{S}{\sigma^2} \left(\mathbf{I}_p - \frac{\mathbf{X}\mathbf{X}'}{W} \right) \cdot \frac{\mathbf{M}_I^*(\boldsymbol{\delta}_\phi)}{p-1} \mid W \right] \right\},$$

$$\eta^*(W) = \frac{1}{g_1(W)} \left(g_3(W) + \frac{1}{\mathbb{E}[(S/\sigma^2)^2 \mid W]} \cdot \text{tr} \left\{ \mathbb{E} \left[\frac{S}{\sigma^2} \cdot \frac{\mathbf{X}\mathbf{X}'}{W} \cdot \mathbf{M}_I^*(\boldsymbol{\delta}_\phi) \mid W \right] \right\} \right).$$

From (17), (32) and Lemma 2, we have

$$\xi^*(W) \leq \frac{1 + W_{p,n}^\xi(\boldsymbol{\delta}_\phi)}{n+p+2} \cdot \frac{\beta_{p,n}^{(2)}}{g_1(W)}, \quad \eta^*(W) \leq \frac{1}{g_1(W)} \left(g_3(W) + \frac{1 + W_{p,n}^\eta(\boldsymbol{\delta}_\phi)}{n+p+2} \cdot \beta_{p,n}^{(2)} \right).$$

Hence by using the same argument as in the proof of Theorem 1, we can complete the proof. \square

G Proof of Theorem 7

For an estimator $\hat{\mathbf{M}}^*(\boldsymbol{\delta}_\phi)$, denote $D(\hat{\mathbf{M}}^*(\boldsymbol{\delta}_\phi)) = \mathbb{E}[L(\hat{\mathbf{M}}^*(\boldsymbol{\delta}_\phi); \mathbf{M}^*(\boldsymbol{\delta}_\phi))]$. Denote $\hat{\mathbf{M}}^*(\xi, \eta; \boldsymbol{\delta}_\phi)$ with $\eta(W) = 1$ by $\hat{\mathbf{M}}^*(\xi; \boldsymbol{\delta}_\phi)$. We first show that when $\xi(\cdot)$ satisfies the conditions of the theorem, $\hat{\mathbf{M}}^*(\xi; \boldsymbol{\delta}_\phi)$ dominates $\hat{\mathbf{M}}_0^*(\boldsymbol{\delta}_\phi)$. In the same way as (31), we have

$$\begin{aligned}
& D(\hat{\mathbf{M}}_0^*(\boldsymbol{\delta}_\phi)) - D(\hat{\mathbf{M}}^*(\xi; \boldsymbol{\delta}_\phi)) \\
&= \mathbb{E} \left[\int_1^\infty \frac{d}{dt} \left\{ (p-1)S^2 g_1^2(W) \xi^2(tW) \right. \right. \\
&\quad \left. \left. - 2Sg_1(W) \xi(tW) \text{tr} [\boldsymbol{\Gamma}'(\mathbf{I}_p - \mathbf{E}_{11}) \boldsymbol{\Gamma} \mathbf{M}^*(\boldsymbol{\delta}_\phi)] \right\} dt \right] \\
&\geq (p-1)\sigma^4 \mathbb{E} \left[\int_1^\infty \frac{d}{dt} \left\{ (S/\sigma^2)^2 g_1^2(W) \xi^2(tW) - 2(S/\sigma^2) g_1(W) \xi(tW) \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right\} dt \right] \\
&= 2(p-1)\sigma^4 \mathbb{E} \left[\int_1^\infty \left\{ (S/\sigma^2)^2 g_1^2(W) \xi(tW) - (S/\sigma^2) g_1(W) \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right\} \xi'(W) dt \right] \\
&= \int \mathbb{E}^v \left[\int_0^{vw} g_1(x/v) f_p(x; \lambda) dx \right. \\
&\quad \left. \times \left\{ \frac{\int_0^w g_1^2(x/v) f_p(x; \lambda) dx}{\int_0^w g_1(x/v) f_p(x; \lambda) dx} \cdot v^2 \xi(w) - v \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right\} \right] \xi'(w) dw \\
&\geq \int \mathbb{E}^v \left[\int_0^w g_1(x/v) f_p(x; \lambda) dz \cdot \left\{ v^2 g_1(w) \xi(w) - v \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right\} \right] \xi'(w) dw. \tag{33}
\end{aligned}$$

The first inequality follows from Lemma 2 and the second inequality follows from the nonincreasingness of $g(\cdot)$. Similar to the argument in the proof of Theorem 2, if $\xi(W)$ satisfy the conditions of the theorem, we have

$$\mathbb{E}^v [v^2 g_1(w) \xi(w) - v \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)] = n(n+2)g_1(w) \xi(w) - n \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \geq 0,$$

and hence $\hat{\mathbf{M}}^*(\xi; \boldsymbol{\delta}_\phi)$ dominates $\hat{\mathbf{M}}_0^*(\boldsymbol{\delta}_\phi)$.

Next we show that when $\eta(\cdot)$ satisfies the conditions of the theorem, $\hat{\mathbf{M}}^*(\xi, \eta; \boldsymbol{\delta}_\phi)$ dominates $\hat{\mathbf{M}}^*(\xi; \boldsymbol{\delta}_\phi)$. In the same way as the above argument, we have

$$\begin{aligned}
& D(\hat{\mathbf{M}}^*(\xi; \boldsymbol{\delta}_\phi)) - D(\hat{\mathbf{M}}^*(\xi, \eta; \boldsymbol{\delta}_\phi)) \\
&= \mathbb{E} \left[\int_1^\infty \frac{d}{dt} \left\{ S^2 g_1^2(W) \eta^2(tW) - 2S^2 g_1(W) \left(g_3(W) + \frac{\phi^2(W)}{W} \right) \eta(tW) \right. \right. \\
&\quad \left. \left. - 2Sg_1(W) \eta(tW) \text{tr} [\boldsymbol{\Gamma}' \mathbf{E}_{11} \boldsymbol{\Gamma} \mathbf{M}^*(\boldsymbol{\delta}_\phi)] \right\} dt \right]
\end{aligned}$$

$$\begin{aligned}
&\geq \sigma^4 \mathbb{E} \left[\int_1^\infty \left\{ (S/\sigma^2)^2 g_1^2(W) \eta^2(tW) - 2(S/\sigma^2)^2 g_1(W) \left(g_3(W) + \frac{\phi^2(W)}{W} \right) \eta(tW) \right. \right. \\
&\quad \left. \left. - 2(S/\sigma^2)^2 g_1(W) \eta(tW) \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right\} dt \right] \\
&= 2\sigma^4 \mathbb{E} \left[\int_1^\infty \left\{ (S/\sigma^2)^2 g_1^2(W) \eta(tW) - (S/\sigma^2)^2 g_1(W) \left(g_3(W) + \frac{\phi^2(W)}{W} \right) \right. \right. \\
&\quad \left. \left. - (S/\sigma^2) g_1(W) \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi) \right\} W \eta'(tW) dt \right] \\
&= 2\sigma^4 \int_0^\infty \mathbb{E}^v \left[\int_0^{vw} \left\{ v^2 g_1^2(z/v) \eta(w) \right. \right. \\
&\quad \left. \left. - g_1^2(z/v) \cdot \frac{v^2 g_3(z/v) + v \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{g_1(z/v)} \right\} f_p(z; \lambda) dz \right] \eta'(w) dw \\
&\geq 2\sigma^4 \int_0^\infty \mathbb{E}^v \left[\int_0^{vw} g_1^2(z/v) f_p(z; \lambda) dz \right. \\
&\quad \left. \times \left\{ v^2 \eta(w) - \frac{v^2 g_3(w) + v \beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{g_1(w)} \right\} \right] \eta'(w) dw \tag{34}
\end{aligned}$$

The first inequality follows from Lemma 2 and the second inequality follows from the nonincreasingness of $g_1(W)$ and the nondecreasingness of $g_3(W)/g_1(W)$. By using the same argument in the proof of Theorem 3, if $\eta(w)$ satisfies

$$\eta(w) \geq \min \left(1, \frac{g_3(w)}{g_1(w)} + \frac{1}{n+2} \cdot \frac{\beta_{p,n}^{(2)}(\boldsymbol{\delta}_\phi)}{g_1(w)} \right),$$

the right hand side of (34) is nonnegative. □

References

- [1] L. C. Adkins and J. B. Eells. Improved estimators of energy models. *Energy Economics*, 17:15–25, 1995.
- [2] A. J. Baranchik. A family of minimax estimators of the mean of a multivariate notomal distribution. *Ann. Math. Statist.*, 41:642–645, 1970.
- [3] G.E. Battese, R.M. Harter, and W.A. Fuller. An error-component model for prediction of county crop areas using survey and satellite data. *J. Amer. Statist. Assoc.*, 83:28–36, 1988.

- [4] R. A. L. Carter, M. S. Srivastava, V. K. Srivastava, and A. Ullah. Unbiased estimation of the mse matrix of stein-rule estimators, confidence ellipsoids, and hypothesis testing. *Econometric Theory*, 6:63–74, 1988.
- [5] G. Casella. Estimators with nondecreasing risk. *Statist. Prob. Lett.*, 10:107–109, 1990.
- [6] B. Efron and C. Morris. Families of minimax estimators of the mean of a multivariate normal distribution. *Ann. Statist.*, 4:11–21, 1976.
- [7] R.E. Fay and R. Herriot. Estimates of income for small places: An application of james-stein procedures to census data. *J. Amer. Statist. Assoc.*, 74:269–277, 1979.
- [8] J. T. Hwang and G. Casella. Minimax confidence sets for the means of a multivariate normal distribution. *Ann. Statist.*, 10:868–881, 1982.
- [9] W. James and C. Stein. Estimation of quadratic loss. In *Forth Berkeley Symp. Math. Statist. Probab.*, volume 1, pages 361–379. University of California Press, Berkley, 1961.
- [10] V. M. Joshi. Inadmissibility of the usual confidence sets for the mean of a multivariate normal population. *Ann. Math. Statist.*, 38:1868–1875, 1967.
- [11] T. Kubokawa. Monotonicity of risk for a shrinkage estimator of a multivariate normal mean. *Comm. Statist. Theory Methods*, 17:499–506, 1988.
- [12] T. Kubokawa. The stein phenomenon in simultaneous estimation: A review. In *Applied Statistical Science, III*, pages 143–173. NOVA Science Publishers, New York, 1998.
- [13] T. Kubokawa and M. S. Srivastava. Estimating risk and the mean squared error matrix in stein estimation. *J. Multivariate Anal.*, 82:39–64, 2002.
- [14] N. G. N. Prasad and J. N. K. Rao. The estimation of the mean squared error of small-area estimators. *J. Amer. Statist. Assoc.*, 85:163–171, 1990.
- [15] C. Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Forth Berkeley Symp. Math. Statist. Probab.*, volume 1, pages 197–206. University of California Press, Berkley, 1956.
- [16] C. Stein. Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.*, 16:155–160, 1964.
- [17] A. T. K. Wan, A. Chaturvedi, and G. Zou. Unbiased estimation of the mse matrices of improved estimators in linear regression. *J. Appl. Statist.*, 30:173–189, 2003.